Parameter and state estimation of nonlinear systems with applications in neuroscience

Michelle Siu Tze Chong

Submitted in total fulfilment of the requirements of the degree of Doctor of Philosophy

March 2013

Department of Electrical and Electronics Engineering

The University of Melbourne

Produced on archival quality paper

Abstract

The focus of this work is deterministic parameter and state estimation of nonlinear systems with applications to neuroscience. Estimation in neuroscience typically involves the reconstruction of unmeasured neural activity from measurements of the human brain. We envisage that estimation plays a crucial role in neuroscience because of the possibility of creating new avenues for neuroscientific studies and for the development of diagnostic, management and treatment tools for diseases such as Epilepsy and Parkinsons disease. One of the most used measurements is the electroencephalogram (EEG). To this end, we consider lumped-parameter nonlinear models with EEG as the output, known as neural mass models.

Four observers are proposed in this thesis: (1) a nonlinear observer, (2) robust circle criterion observers, (3) an adaptive observer and (4) the supervisory observer. These observers are synthesised for classes of nonlinear systems, that cover some of the commonly used neural mass models. Two state observers are shown and designed respectively, in Part I, to be robust towards input and measurement noise, as well as small perturbations in parameters. In the absence of noise and perturbations, the estimates converge exponentially to the true values. The convergence of estimates to their true values is with some error in the presence of noise and perturbations. Chapter 3 presents a nonlinear observer specific to the class of neural mass models considered. In Chapter 4, we propose robust circle criterion observers for a class of systems, that covers all our examples. We extended available results in the literature such that they can be synthesised for the neural mass models. The robustness of the designed state observers towards parameter uncertainty motivates the estimation of both parameters and states in Part II.

In Chapter 5, we design an adaptive observer for a class of interconnected neural mass models. The convergence of the estimates is asymptotic. Finally, in Chapter 6, we present an alternative method using a multiple-model architecture, known in the literature as the supervisory framework. Under non-restrictive conditions, we guarantee the practical convergence of parameters and states.

Declaration

This is to certify that

1. the thesis comprises only my original work towards the PhD except where indicated in the Preface,
2. due acknowledgement has been made in the text to all other material used,
3. this thesis is fewer than 100,000 words in length, exclusive of tables, maps, bibliographies and appendicies.

Michelle S. Chong
Date: 26 March 2013

Preface

The material developed in this thesis is the result of original research, unless otherwise stated, solely during the author's candidature in the Department of Electrical and Electronic Engineering, the University of Melbourne under the supervision of Prof. Dragan Nešić and Dr. Levin Kuhlmann. The results presented in this thesis have been obtained in collaboration with the author's supervisors and Dr. Romain Postoyan from Centre de Recherche en Automatique de Nancy (CRAN), France. Some results were also obtained in collaboration with Dr. Andrea Varsavsky, who was also a supervisor when she was in the University of Melbourne.

The work presented in this thesis have been accepted or submitted for publication in journals, as well as presented and accepted at international conferences:

1. **M. Chong**, R. Postoyan, D. Nešić, L. Kuhlmann and A. Varsavsky. 'A robust circle criterion observer with application to neural mass models' in Automatica, vol. 48, pp 2986-2989, 2012.

2. **M. Chong**, R. Postoyan, D. Nešić, L. Kuhlmann and A. Varsavsky. 'Estimating the unmeasured membrane potential of neuronal populations using a class of deterministic nonlinear filters' in the Journal of Neural Engineering, vol. 9., pp 026001, 2012.

3. **M. Chong**, D. Nešić, R. Postoyan and L. Kuhlmann. 'Parameter and state estimation of nonlinear systems using a multi-observer under the supervisory framework' - *submitted to IEEE Transactions of Automatic Control.*

4. **M. Chong**, R. Postoyan, D. Nešić, L. Kuhlmann and A. Varsavsky. 'A nonlinear estimator for the activity of neuronal populations in the hippocampus' in the Proceedings of 18th World Congress of the International Federation of Automatic Control (IFAC), 2011.

5. **M. Chong**, R. Postoyan, D. Nešić, L. Kuhlmann and A. Varsavsky. 'A circle criterion observer for estimating the unmeasured membrane potential of neuronal populations' in the Proceedings of the Australian Control Conference, 2011.

6. R. Postoyan, **M. Chong**, D. Nešić and L. Kuhlmann. 'Parameter and state estimation for a class of neural mass models' in the Proceedings of the 51st IEEE Conference on Decision and Control, 2012.

Acknowledgements

I would like to take this opportunity to express my gratitude to the many people who have provided help and encouragement in the period leading up to and during the development of this thesis.

I am tremendously lucky to have had the opportunity to work with Dragan Nešić, Levin Kuhlmann, Romain Postoyan and Andrea Varsavsky on the ideas in this thesis. First and foremost, I would like to thank my thesis supervisor, Professor Dragan Nešić for his guidance, insight, support, harsh (oops!) encouragement and for agreeing to take me on as a graduate student. I cannot thank him enough for the generosity of his time and experience. He is my first teacher in nonlinear control theory and his commitment to research has sparked a committed interest in me to pursue a research career in mathematical control theory and its applications.

I would also like to thank my thesis co-supervisors, Dr. Levin Kuhlmann and Dr. Andrea Varsavsky, who provided their knowledge in matters related to neuroscience and epilepsy, in particular. Levin introduced me epilepsy research as an undergraduate under the summer research scholarship programme by the School of Engineering, the University of Melbourne and later, as a student in the Bio21 Cluster Undergraduate Research Opportunities Program (UROP). This motivated the development of observers presented in this thesis for the epileptic seizure detection effort. My thesis supervisors have continuously supported me with their respective knowledge in control theory and neuroscience, and by their efficiency in dealing with practical issues such as funding and bureaucracy.

I am very grateful to Dr. Romain Postoyan for first initiating me to the realm of deterministic nonlinear observer design and for all our valuable discussions that lead to the results presented in this thesis. I would like to thank him for his warm hospitality in my two visits to Europe during the course of my PhD studies by welcoming me to Centre de Recherche en Automatique de Nancy, France. His constant flow of ideas is an inspiration to a beginning research student and his relentless pursuit for perfection is reflected in all the papers we have co-authored.

Thanks also to A/Prof. Leigh Johnston who served on my research higher degree committee, for her time and effort in attending my seminars and progress meetings.

I would like to acknowledge the Epilepsy research group with members from St. Vincent's Hospital (Melbourne), Bionics Institute and the Department of Electrical and Electronic Engineering, the University of Melbourne for the introduction to the many facets of epilepsy research in an interdisciplinary environment. Thanks to the

members who took the time to explain their work to me as an undergraduate and for the initial discussions at the start of my graduate studies: Anthony Burkitt, Mark Cook, Karen Fuller, Dean Freestone, David Grayden, Colin Hales, Alan Lai, Stephan Lau, Elma O'Sullivan-Greene, Andre Peterson and Simon Vogrin.

I am grateful to my teachers over the course of my graduate studies, Prof. Jonathan Manton, A/Prof Michael Cantoni and Prof. Subhrakanti Dey for providing a solid primer to convex optimisation, linear systems theory and statistical signal processing, respectively. I am also indebted to the instructors at the Elgersburg School, Prof. Randy Freeman for nonlinear control and Prof. Laurent Praly for nonlinear observer design. The discussions in class, during meals and even during games of German bowling, enhanced my understanding in nonlinear control theory and inspired me.

PhD scholarships from the Department of Electrical and Electronic Engineering, the University of Melbourne as well as the Research Endowment Fund, St. Vincent's Hospital (Melbourne) have supported my research. I am also thankful for the Melbourne Abroad Travelling Scholarship, the M.A. Bannett Research Scholarship Special Travel Grant-in-aid, the Melbourne School of Engineering Research Training Conference Assistance Scheme, a travelling scholarship from the Elgersburg School and the grants of my supervisors for supporting my travels.

On a personal note, I would like to thank my friends and colleagues at the department who made my graduate life enjoyable and whom I have yet to mention. Thanks to Adel Ali, Mathias Foo, Rika Hagihara, Anh Trong Huynh, Tatiana Kameneva, Sachintha Karunaratne, Robert Kerr, Sei Zhen Khong, Kelvin Layton, Merid Ljesnjanin, Juan Lu, Amir Neshastehriz, Emily O'Brien, Sajeeb Saha, Senaka Samarasekera, Laven Soltanian, Martin Spencer, Bahman Tahayori, Tuyet Vu, Mohd Asyraf Zulkifley, for their friendship and support. I would also like to thank my friends outside the lab, of which there are far too many to mention. Special thanks to Khai Li Chai, Amelia Chia, Melissa Chong, Yi Man Lee, Samuel Loh, Su-Min Wong and Michelle Yew.

Last but not least, I am forever indebted to my family, Roger, Mary, Crystal and Charmaine, for their love and support. My parents have made countless sacrifices for me and are always supportive in my choices, especially at important cross-roads in my education. I would also like to thank Uncle Lee and family for their kindness and support. This thesis is dedicated to them and in the fond memory of my dad, who instilled the love of learning in me and who left us too soon.

Michelle S. Chong

Melbourne, Australia. February 2013.

Contents

List of Figures

List of Notations

$\mathbb{R}$	The set of real numbers.
$\mathbb{R}_{\geq 0\,(>0)}$	The set of non-negative (strictly positive) real numbers.
$\mathbb{R}^{n\times m}$	The space of real matrices with dimensions $n \times m$.
(a, b)	A vector $\begin{bmatrix} a \\ b \end{bmatrix}$, for all $a \in \mathbb{R}^{n_a}$ and $b \in \mathbb{R}^{n_b}$.
diag	The matrix $\begin{bmatrix} A_1 & 0 \dots & 0 & \\ 0 & \dots & & \vdots \\ \vdots & & \dots & 0 \\ 0 & \dots & 0 & A_n \end{bmatrix}$, where A_i for $i \in \{1, \dots, n\}$ are $m \times m$ matrices, is denoted as $\mathrm{diag}(A_1, A_2, \dots, A_n)$.
$\lambda_{\max}(P)$ $(\lambda_{\min}(P))$	The maximum (minimum) eigenvalue of a real, symmetric matrix P.
$\star$	The symmetric block component of a symmetric matrix.
$\mathbb{I}$	The identity matrix.

$\lvert f(t)\rvert$	The vector norm of f at each time t.
$\lVert f(t)\rVert_2$	The $\mathcal{L}_2$ norm of f.
$\lceil a \rceil$	The smallest integer larger than or equal to $a \in \mathbb{R}$.
$\eta \in \mathcal{N}(\mu, \sigma^2)$	A signal $\eta \in \mathbb{R}$ that is drawn from a Gaussian distribution with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 \in \mathbb{R}$.
B_r	The closed ball centered at 0 for some $r > 0$ is denoted by $B_r := \{s \in \mathbb{R}^n : \lvert x\rvert \leq r\}$.
$\mathcal{L}_\infty$	The set of functions $f : \mathbb{R} \to \mathbb{R}^n$, for some $n \in \mathbb{Z}$, such that for any $0 \leq t_1 \leq t_2 \leq \infty$ there exists $r \geq 0$ so that $\lVert f\rVert_{[t_1,t_2]} := \sup_{\tau \in [t_1,t_2]} \lvert f(\tau)\rvert \leq r$.
$\mathcal{M}_\Delta$	The set of piecewise continuous functions from $[0, \infty)$ to B_Δ, for any $\Delta > 0$.
$x_{[t_1,t_2]}$	For a signal $x : \mathbb{R}_{\geq 0} \to \mathbb{R}^n$, $x_{[t_1,t_2]}$ is the signal $x(t)$ considered on the interval $t \in [t_1, t_2]$.
$(\cdot)^-$	The left-limit operator.
$\mathcal{K}$	A continuous function $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is said to be a class $\mathcal{K}$ function, if it is strictly increasing and $\alpha(0) = 0$.
$\mathcal{KL}$	A continuous function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is said to be a class $\mathcal{KL}$ function, if, for each fixed s, the mapping $\beta(r, s)$ is a class $\mathcal{K}$ function with respect to r and for each fixed r, the mapping $\beta(r, s)$ is decreasing with respect to s and $\beta(r, s) \to 0$ as $s \to \infty$.

Chapter 1

Introduction

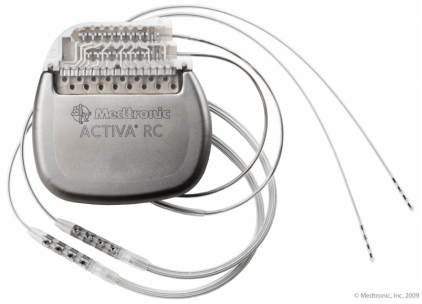
1.1 Motivation and scope

Approximately 50 million individuals in the world have fits or convulsions, known as seizures caused by neurological disorders such as epilepsy and Parkinson's disease [5]. These individuals face the uncertainty of uncontrolled seizures that restricts their daily activities, such as swimming, cooking and driving. In some parts of the world, they face social stigma and exclusion that ranges from misunderstanding, disadvantages in employment to serious consequences brought on by legislation. In fact, up till 1956, 18 states in the United States of America provided eugenic sterilisation of epileptic patients [3]. The number of individuals *affected* by epilepsy is even higher, at around 200 million when taking into account the family members and friends of the patient [4]. The cost of epilepsy is therefore not only physical, but also social and psychological.

The goal of epilepsy treatment is prevention rather than cure. Fortunately, 70 % of patients can become seizure free with proper medication. The remainder require resective surgery or implantable devices for seizure control. Figure 1.1 shows some implantable devices which have been approved for clinical research only in the United States [1; 2]. The Deep Brain Stimulation (DBS) device by Medtronic Inc. is still undergoing clinical trial [1]. Some of these devices monitor electrical brain activity in a localised region for seizure related activity and delivers controlled electrical pulses to the region believed to be the seizure foci. The algorithm that monitors brain activity for abnormality is known in the literature as a *seizure detection/prediction algorithm* and the measurement of electrical activity is the electroencephalogram (EEG). The high false positive rates of these algorithms render these devices undesirable for clinical

use and the quest for a successful device continues [137].

(a) An epileptic seizure control device by NeuroPace Inc., Mountain View, California. (Credit: NeuroPace Inc.)



(b) A Deep Brain Stimulation (DBS) device first developed for abnormal activity caused by Parkinson's disease by Medtronic Inc., Minneapolis, Minnesota. (Credit: Medtronic Inc.)

Figure 1.1: Implantable devices

Seizure detection/prediction algorithms are predominantly non-model based, e.g. [56; 93; 106; 107; 117]. Non-model based algorithms extract features from the EEG which are used to classify neural phenomena into seizure and non-seizure activity. Usual features are temporal, spectral or both. These methods are continuously under development [109].

We believe that a model-based approach is needed to produce a successful seizure detection/prediction algorithm. By model-based, we mean the usage of mathematical models in estimating neural activity from measurements such as the electroencephalogram (EEG). Model-based estimation has proved to be successful in many other application areas such as marine engineering for the positioning of ships in the ocean, robotics, ecology for the study of food-chain system and many more [113]. The model-based approach is starting to gain ground in neuroscience [124]. We envisage that estimation is essential for seizure detection/prediction due to seizure indicators that differ from patient to patient. Since useful models that can capture seizure activity are nonlinear, recent advances in nonlinear mathematical control theory may provide the much needed backing in bridging this gap. This thesis is a step in this direction by rigorously designing provable *model-based estimation algorithms*, also known as **observers** tailored for this endeavour (Figure 1.2).

The human brain can be viewed as a system with inputs and outputs or measure-

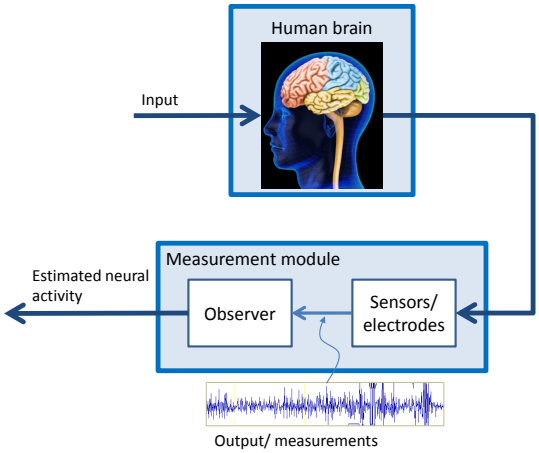


Figure 1.2: Model-based estimation of neural activity (observers)

ments. Examples of measurements include the electroencephalogram (EEG), functional magnetic resonance imaging (fMRI) and magnetoencephalogram (MEG), to name a few. We focus on using the EEG as a measurement due to its cost-effectiveness and portability such that it is usable in implantable devices. The EEG also has a very high temporal resolution, in the order of milliseconds, a characteristic that is useful when detecting high frequency oscillations (HFOs) with frequencies greater than 80 Hz which are postulated to be an indication of seizure activity [136; 145; 164].

Due to limitations in technology, the EEG is unable to capture all neural activities of interest. The estimation of unmeasured neural activities is therefore very appealing to clinicians and neuroscientists. This is due to the possibility of opening up new avenues for neuroscientific studies, diagnostics and the treatment of neurological disorders via the development of monitoring or control strategies [69].

One of the primary advantages of estimation lies in reducing the number of sensors/electrodes needed and thereby keeping surgical invasion to a minimum. This is particularly useful for neurological events such as seizures caused by epilepsy or pathological dynamics of Parkinson's disease that are generated in deep brain structures such as the hippocampus and the basal ganglia, respectively. Implanting the smallest required number of electrodes minimises risks such as infection and haemorrhage as well as pain that results from implantation.

We provide several examples where an observer is used in the field of neuroscience.

- **Neuroscientific studies**

  The observer serves to enhance neuroscientific studies. The estimation of unmea-

sured neural activities in the setup shown in Figure 1.2 provides neuroscientists with a basis for generating experimentally provable hypotheses on the underlying mechanisms that govern a neurological event, such as seizures caused by epilepsy or Parkinson's disease. Works by Tokuda et al. in [143], Totoki et al. in [144], Tyukin et al. in [148] and Mao et al. in [103] perform model-based estimation at the single neuron level for various neuroscientific purposes. In [150], Ullah and Schiff performed model-based estimation of the ion concentrations of hippocampal neurons from membrane potential measurements of a neuron to study the dynamics of ion concentrations during epileptic seizures. Advances in this area will fuel the development of diagnosing, monitoring and treatment strategies for neurological diseases as discussed in the sequel.

- **A diagnostic system**
  Experimental studies have provided clinicians and neuroscientists clues to the presence of certain indicators that suggest the possibility of epilepsy in patients [136]. An observer would provide estimates of indicators from measurements. For example, the synaptic gains of the pyramidal neurons, the excitatory and inhibitory populations in the CA1 region of the hippocampus have been identified in [155] as parameters that are related to seizure and non-seizure activity. Clinicians may be aided by estimates provided by an observer to diagnose patients whose parameter estimates consistently belong to the seizure-related range.

- **A seizure warning system**
  This takes the form of using estimates of neural activity from an observer to discern seizure from non-seizure behaviour by the classifier in the seizure prediction module (see Figure 1.3). To this end, key features of a seizure need to be known. Work has been done in identifying parameters of models that are postulated to be related to seizure activity [153]. We discuss this model in greater detail in Chapter 2. It is the task of the classifier to raise a red flag when the parameters fall within the range that have been predetermined to be seizure-related.

- **Closed-loop seizure control system**
  The open-loop warning system can then initiate the administration of drugs or electric stimulation to abate seizures, which we call a closed-loop control system, see Figure 1.4. The observer provides estimates of neural activity such that a control law may be formulated to trigger seizure suppression procedures. We discuss the design of a **control law** (see Figure 1.4) in Chapter 7.1, where we present future directions of this work.

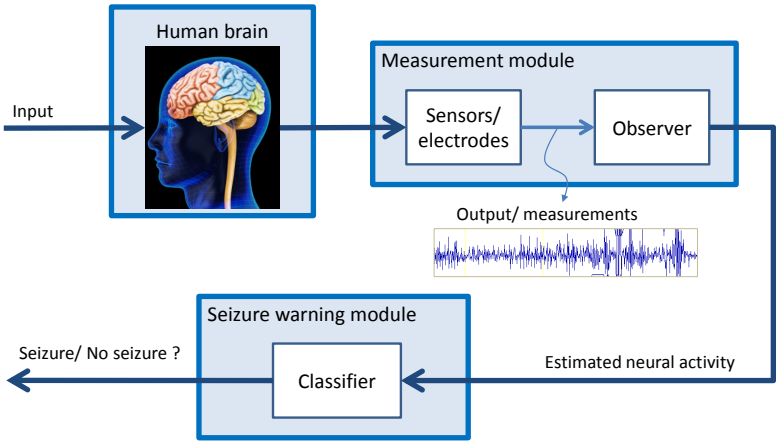


Figure 1.3: An open-loop warning system

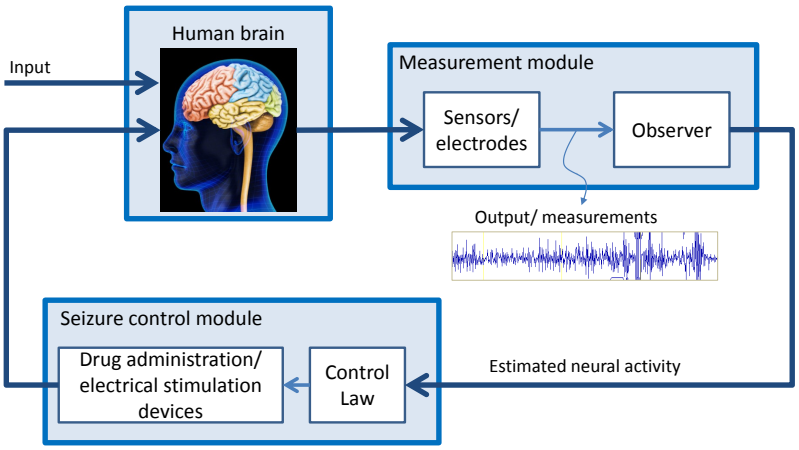


Figure 1.4: Closed-loop seizure control system

The design of observers in the examples mentioned above requires the use of mathematical models that appropriately describe the part of the brain concerned and the measurement used. We briefly discuss our choice of neural models in the following section.

1.2 Neural model

As mentioned earlier, the EEG is our measurement of choice. Here, we have used the term EEG as the umbrella term that includes measurements obtained from the scalp

or from the surface of the cortex, sometimes known in the literature as the electrocorticography (ECoG) or intracranial EEG (iEEG). The sensors/electrodes used to obtain the EEG are implanted on the scalp or surface of the brain. Each sensor provides a single channel of EEG which is proportional to the electrical discharges of a *population* of neurons in a localised volume perpendicular to the surface [115]. This volume of the brain is known as a *cortical column* [111, Chapter 7], see Figure 1.5.

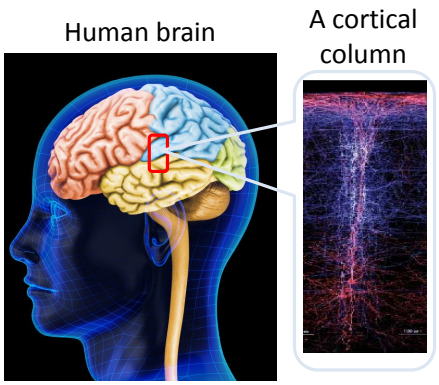


Figure 1.5: A cortical column. (Diagram credit: iStockphoto/Guido Vrola and http://www.mada.org.il)

Within a cortical column, the neurons may be categorised by function into inhibitory and excitatory neurons. They can be pyramidal and basket cells in the hippocampus or pyramidal cells and thalamic interneurons in the cortico-thalamic system [127]. In 1973, Wilson and Cowan suggested a model for the interaction between inhibitory and excitatory populations of neurons in [156]. This model forms the basis of many models, including the neural mass models which we will discuss in greater detail in Chapter 2.

We view each cortical column as a dynamical system Σ (see Figure 1.6), with measurements or outputs y and inputs u. Examples of an output from a cortical column model are the average electrical activity of neural populations in the cortical column or the EEG measurement. Examples of an input are the electrical activity from neighbouring cortical columns or electrical pulses sent from an implantable device to suppress seizures. Each system Σ has internal variables or states x and is parameterised by p^*, which forms a family of parameterised systems (Figure 1.6).

In general, depending on the measurement y considered, the states x may capture the membrane potential of a neuron (microscopic) or the mean membrane potential of neuronal populations (mesoscopic) and the parameters p^* may be the neural membrane

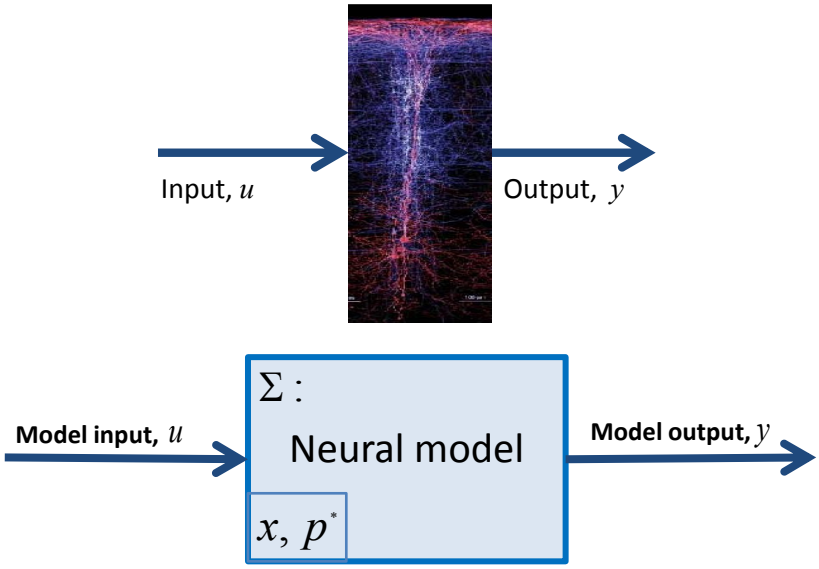


Figure 1.6: A cortical column is a system whose dynamics are described by a mathematical model.

capacitive current (microscopic) or the synaptic gain of populations (mesoscopic). The states are physiologically relevant and often unmeasurable by conventional means. The parameters p^* are identified to vary as a result of different brain phenomena and hence usually resides in a known, compact set Θ. For instance, in the case of a model of epileptic activity in the hippocampus [153], the synaptic gains of the neuronal populations were identified to be the parameters that change when the brain transitions from normal to epileptic activity. Hence, the possibility of detecting such changes and to then develop seizure control strategies to mitigate seizures form great motivation for the estimation of states x and parameters p^*.

In this thesis, our measurement of choice is the EEG, which is of the mesocopic scale. Hence, we consider cortical column models that are lumped-parameter models, also known in the literature as neural mass models, a term coined in [39]. These models have the added advantage of being able to capture the desired neural phenomenon such as seizures, while still being of a low enough dimension to be useful in the estimation and control theoretic sense. This class of systems is represented by ordinary differential

equations and they can be written in state space form as follows:

$$\begin{aligned} \Sigma: \qquad \dot{x} &= f(x, u, p^*) \\ y &= h(x, p^*), \end{aligned} \tag{1.1}$$

where the state is $x \in \mathbb{R}^{n_x}$, the input is $u \in \mathbb{R}^{n_u}$, the parameter is $p^* \in \Theta \subset \mathbb{R}^{n_p}$, measurement/output is $y \in \mathbb{R}^{n_y}$ and n_x, n_u, n_p and n_y are positive integers. Examples of commonly used neural mass models considered in this thesis are: models for the generation of alpha rhythms by Jansen and Rit as well as Stam et. al., respectively in [74; 138] and a model of epileptic activity in the hippocampus by Wendling et. al. in [155]. These models were based on the seminal work by Wilson and Cowan [156], Freeman [50] and Lopes da Silva et. al. [99; 100]. We describe these models in greater detail in Chapter 2.

1.3 The estimation problem

The objective is to estimate the states x and parameters p^* of the model Σ. The dynamical system that computes these estimates is known as an '*observer*' or '*estimator*' or '*filter*' [12; 23]. In this thesis, we will use the term 'observer'. When the parameter p^* is known, we design a **state observer** to reconstruct the states x of the model Σ. The observer Σ_o takes in the available information (the input u and measurement y) to provide state estimates $\hat{x}$. When the parameter p^* is unknown, Σ_o is called a **state and parameter observer**, which provides estimates of both states $\hat{x}$ and parameters $\hat{p}$. Figure 1.7 illustrates this setup.

The task of observer design can be approached in two ways: stochastic or deterministic. The stochastic approach to estimation is more popularly known in the literature as 'filtering' and the system is modelled by stochastic differential/difference equations [75]. Classical, statistical methods such as least-squares, maximum-likelihood and minimum variance estimation were first applied to linear filtering problems and later extended to nonlinear cases [83]. A popular point of view is Bayesian-based, i.e. the estimate of the state can be constructed from the conditional probability density function of the state, given the available measurements. See [75] for a unified treatment of bayesian-based linear and nonlinear filtering. Popular stochastic observers include the Kalman filter and its variants[1]. These filters are used in many applications, including neuro-

[1]The Kalman filter can also be formulated deterministically, where its design is reducible by duality to a linear quadratic optimal control problem [132, Section 8.3].

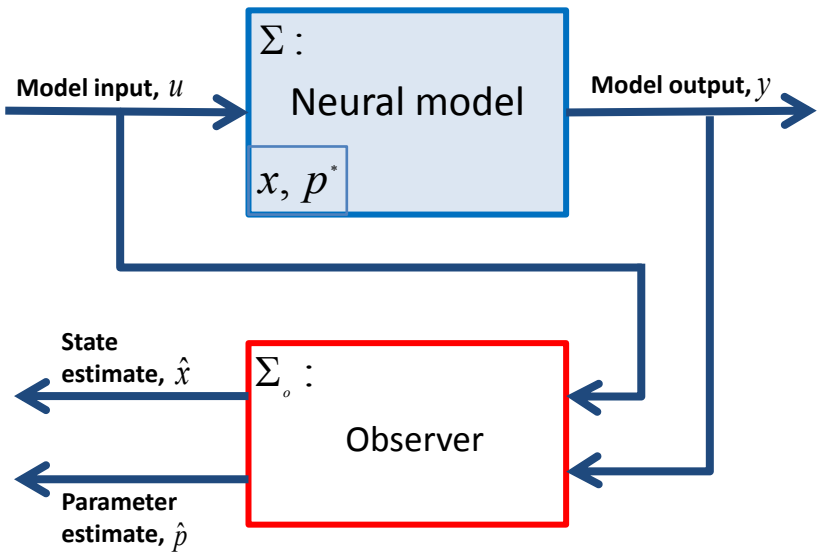


Figure 1.7: Observer setup

science [51; 52; 125; 150]. One popular usage of filters in neuroscience is a framework known as Dynamic Causal Modelling [37; 88; 140], where estimates from the filters are used to choose the best model that describes the measurements under some conditions. Stochastic methods provide convergence of estimates with some probability. In other words, the convergence of the estimates to the true values is not guaranteed for every trajectory.

On the other hand, the deterministic approach guarantees convergence of estimates. For the simpler problem of state estimation only, the case for linear systems is completely solved in the form of a Luenberger observer [101] and the deterministic Kalman and Bucy filter [78]. This is not the case for nonlinear systems, where no general solution exists. The design of nonlinear observers is done on a 'case-by-case' basis, which is reliant on the mathematical structure of the model. Hence, it makes sense to consider a class of systems that share the same mathematical structure and an observer is designed for that class of systems. Due to the widespread applicability of nonlinear observers, nonlinear state observer design has been extensively studied [6; 7; 13; 15; 16; 17; 18; 19; 23; 26; 27; 28; 29; 40; 43; 44; 45; 47; 48; 53; 54; 55; 59; 60; 61; 70; 81; 82; 84; 87; 90; 91; 94; 97; 101; 102; 104; 108; 113; 119; 120; 121; 123; 126; 128; 129; 130; 142; 146; 147; 157; 158; 159; 161; 163]. Deterministic state observers can generally be categorised into observers for special classes of nonlinearities (e.g. Lips-

chitz nonlinearities [54; 121], monotone nonlinearities [16; 158]) and high gain observers [13; 87].

We contribute to this body of work with the design of two state observers in Chapters 3 and 4 for a class of nonlinear systems which includes some of the neural models introduced earlier: a Luenberger-like observer and a robust form of the circle criterion observers, first introduced in [16]. The existence of a circle criterion observer is based on the feasibility of a linear matrix inequality (LMI). Existing circle criterion observers in [16; 17; 45; 158] do not yield LMIs that are feasible for the neural mass models we consider. Hence, we combined several ideas from these papers including considering globally Lipschitz nonlinearities and introducing a multiplier to obtain LMIs that are feasible for our examples. We further improved the design by allowing the user to specify attenuation factors towards noise.

Similarly for the case of parameter *and* state estimation. One common approach is to augment the state vector with the parameters and employ a state observer to estimate the augmented vector such that we obtain both state and parameter estimates. This approach is not conducive even for the case of linear systems because doing so may turn the augmented system highly nonlinear, where the synthesis of a nonlinear observer is difficult.

An alternative is the design of adaptive observers [20; 22; 26; 32; 46; 57; 89; 104; 105; 122; 139; 149; 158; 160; 161]. An adaptive observer may be viewed as a state observer with an adaptive law that provides parameter estimates jointly to obtain state estimates. We will take this approach in Chapter 5 where we design an adaptive observer for a class of nonlinear systems which includes the interesting case of interconnected neural mass models. This adaptive observer drew inspiration from the design in [46] which combines the high gain idea used in state observers [13; 87] with an adaptive observer design for linear time varying systems by Zhang and Clavel in [162].

Another approach uses the multiple-model architecture (see [9] and references therein for an overview). This architecture employs a bank of state observers, where each observer is designed for a particular nominal parameter value chosen from a known set, to provide state and parameter estimates under some scheme. It has traditionally been pursued using *stochastic* methods [12, Section 8.4] and has recently been studied in the *deterministic* sense [10], [11]. Encouraged by recent results in supervisory control (see [64; 152] for linear systems and [21] for nonlinear systems), which uses the multiple model architecture for stabilisation, we adapted it for estimation purposes in Chapter 6. We call this adapted setup, a supervisory observer and guarantee the convergence of states and parameter estimates in the deterministic sense under certain conditions.

Two main issues arise in the estimation problem: Firstly, how 'good' are the estimates $\hat{x}$ and $\hat{p}$? Secondly, how much modelling error, measurement and input noise can the designed observers tolerate? We will be addressing these questions in this thesis.

1.4 Thesis contributions and outline

The focus of this thesis is the design of **deterministic state and parameter observers** for nonlinear systems tailored for applications in neuroscience. Nevertheless, we stress that our results also apply to any other models in other applications that fit the classes of systems we consider that satisfy our assumptions. The results of this thesis have been published in several international peer-reviewed publications and the references are listed in the Preface. We summarise our contributions:

- To the best of our knowledge, all the observers designed in this thesis are the first for the class of neural mass models presented in Chapter 2. The deterministic convergence of the estimates are rigorously proved using tools from nonlinear stability theory.
- Estimation issues faced in neuroscientific studies are also taken into account in our designs of the state observers, i.e. the presence of disturbances and uncertainties. We analyse the robustness of the state observers with respect to parameter, model and input uncertainty as well as measurement disturbance. One particular design allows the user to attenuate the influence of these uncertainties and disturbances. Several application scenarios are simulated to illustrate these properties.
- A framework for parameter and state estimation using *state* observers is adapted from supervisory control and rigorously proven to provide practical convergence of estimates for nonlinear systems with *general* structure under some conditions. We show that these conditions are not restrictive by applying our general results to linear systems and to a class of nonlinear systems that includes the neural mass models considered in Chapter 2.

We now give a brief outline for the material that is developed in the subsequent chapters:

Chapter 2 introduces the class of neural mass models we consider. This class of models shares the same mathematical structure that makes them highly amenable from an estimation and control viewpoint. Three representative models are described, they are:

(i) a model by Stam et. al. in [138] that described alpha rhythms seen in the EEG, (ii) a model by Jansen and Rit in [74] that describes the generation of alpha rhythms in the cerebral cortex and (iii) a model by Wendling et. al. in [153] that describes epileptic activity in the hippocampus. We will provide a brief overview of the physiological interpretation of these models and then present them in state space form, which is most convenient for observer design.

This thesis can be read in two parts: **Part I** (Chapters 3-4): State estimation and **Part II** (Chapters 5-6): Parameter and state estimation.

Part I: State estimation

Chapter 3 presents a nonlinear observer that is designed for this class of neural mass models. Some highly desirable features of these models allow us to design an estimator that has a state error system with a cascaded structure. This allows us to apply the interesting result of ISS for cascaded systems in stability analysis [86, Lemma 4.7]. We will also show that the designed observer has desirable robustness properties.

Chapter 4 introduces a robust circle criterion observer that we have designed. This observer takes into account two main robustness issues encountered in neuroscientific studies, that is input uncertainty and measurement noise. We allow the user to specify attenuation factors towards these undesirables and should a derived LMI be solvable, a robust circle criterion observer can be obtained.

The success of state estimation and the robustness of the designed observers toward small perturbations in parameters motivated the next step towards achieving our goal, that is parameter and state estimation in the following part.

Part II: Parameter and state estimation

Chapter 5 presents an adaptive observer for the class of neural mass models we consider, which can be written as a linear part and a triangular nonlinear part that is linearly parameterised. We exploited this structure to design an adaptive observer that is applicable to a subset of the class of models considered and most interestingly, interconnected models in this subset.

Chapter 6 proposes using a multi-observer to provide state and parameter estimates for *general* nonlinear systems under the supervisory framework. We call this setup the

supervisory observer. The results obtained are applicable to *general* nonlinear systems and we show that our main results can be applied to the class of neural mass models considered in Chapter 2.

Finally, Chapter 7 concludes this thesis with some discussion for future work. Appendix A serves as a primer on the stability tools used in observer design and the analysis of systems in this thesis. Appendix B contains all mathematical proofs of results and lastly, Appendix C lists standard values of the constants used in the considered class of neural mass models.

Chapter 2

Neural models

We first provide a definition of the basic neurophysiological terminology used in the literature and in this thesis. The intention is not to be comprehensive, but serves to set the stage for the types of neural models considered for observer design and the physiological meaning of the states and parameters estimated. We then introduce the class of neural models considered in this thesis and write these models in state space form in state coordinates that are convenient for observer design.

2.1 Basic neurophysiological and electroencephalographical terminology

The respective definitions of the neurophysiologically related terminology below can be found in [38; 79; 127] and electroencephalography related terminology can be found in [115]. Here, we present a glossary of basic terms used.

- *Neurons* are cells that are found in the brain. There is an estimate of more than 10^{11} neurons in the human brain.
- A neuron consists of a *cell body* or *soma*, *dendrites* and *axon*. See Figure 2.1.
- *Action potentials* or *impulses* or *spikes* are generated by neurons and they are the communication signal between neurons. These signals are received by dendrites, processed in the soma and the neuron outputs signals to other neurons via its axon.
- The neuronal response to stimuli is often a sequence of *action potentials* which can be characterised through their timing. The *firing rate* captures the average

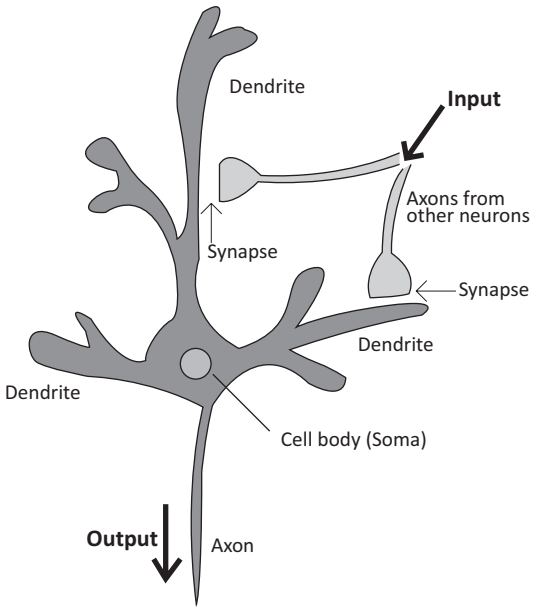


Figure 2.1: A neuron. (Diagram credit: Crystal Chong)

number of spikes (*action potentials*) in a time interval.

- The *synapse* is the junction between the dendrite of one neuron and the axon of another neuron. There are *electrical* and *chemical* synapses.
- *Electrical synapses* are direct, electrically conductive junctions.
- *Chemical synapses* transmit signals from the *pre-synaptic* cells to the *post-synaptic* cells that are separated by the *synaptic cleft*. When an action potential is received by an axon, chemical and electrical reactions are triggered which causes the release of *neurotransmitters* into the synaptic cleft. They diffuse across the synaptic cleft and react with the *transmitter-gated ion channels* (i.e. *receptors*) on the postsynaptic cell, causing a change in the *postsynaptic potential* (PSP).
- The *postsynaptic potential* (PSP) is in the milli-Volts (mV) range.
- The *synaptic gain* or *strength* of a single synapse is quantified by the gain in the amplitude of the PSP as a result of a pre-synaptic action potential. The synaptic gain of a single synapse is proportional to the amount of neurotransmitter released and the number of postsynaptic receptors. The total synaptic gain from the presynaptic cell to the postsynaptic cell is proportional to the number of connections from the presynaptic cell to the postsynaptic cell.

- A neuron can either be *excitatory* or *inhibitory*. An *excitatory* neuron transmits action potentials to a receiving neuron causing an increase in PSP in the receiving neuron. This event is known as *depolarisation* and is amenable to the generation of action potentials. On the other hand, a neuron that causes a decrease in PSP or *hyperpolarisation* is an *inhibitory* neuron. This decreases the likelihood of action potential generation. Excitation and inhibition are mediated by multiple neurotransmitters. An example of a major neurotransmitter is the *γ-aminobutyric acid (GABA)*, which is commonly found in the *hippocampus*.

- The *hippocampus* is one of the most well-studied parts of the mammalian brain. Most epileptic patients have seizures that involve the hippocampus. The principal neurons in the hippocampus are the *pyramidal neurons*. Depending on their size and appearance, the pyramidal cell layer is divided into three regions labelled CA1, CA2 and CA3. A member of the class of neural mass models considered (the model by Wendling et. al. in [155]) models the CA1 region.

- *Intrinsic neurons* or *interneurons* are a type of neuron with a locally restricted axon plexus that lack spines and release γ-aminobutyric acid (GABA). There are two types of GABA dynamics in the CA1 pyramidal neurons: a fast response near the soma and a slow response near the dendrites. These populations are included in the model by Wendling et. al. in [155].

- A *neural population* refers to a group of neurons characterised by location (e.g. cortex, hippocampus) or by type (e.g. excitatory, inhibitory, pyramidal). Common nomenclatures include a *cortical column*, *excitatory population* or *inhibitory population*. The mean PSP or membrane potential of a neural population is the average potential of all the neurons in that population.

- A *cortical column* is a group of neurons in the cortex that reside in a column of $300-500\mu m$ in diameter perpendicular to the surface of the cortex [111, Chapter 7].

- The *electroencephalogram* (EEG) measures the electrical activity of neural populations [115]. It is said to reflect the mean PSP or mean membrane potential of neural populations. The sensors used to measure EEG are electrodes that can be placed on the scalp and the measurement is known as '*scalp EEG*' or near the surface of the brain to measure '*intracranial EEG (I-EEG)*', '*electrocorticography (ECoG)*' or '*subdural EEG (SD-EEG)*'. Often in epilepsy studies, *depth* electrodes

are inserted in brain structures such as the hippocampus to capture the focus of the seizure, that is otherwise poorly localised in the scalp EEG.

- *Alpha rhythms* are oscillations of $9-11$ Hz observed in the EEG. These patterns are often recorded when the subject's eyes are closed in a relaxed, awake state.

2.2 Neural models

Mathematical models in neuroscience can be largely classified into comprehensive models and heuristic models. Comprehensive models are constructed by taking into account all known neurophysiological facts and data. An example is the Nobel-prize winning model of the initiation and propagation of action potentials in a single neuron by Hodgkin and Huxley [68]. As the facts and data considered increases (e.g. modelling a *population* of neurons as opposed to a single neuron), so does the complexity of the model increase. Since neuroscience is a largely evolving field, we are still far from painting a complete picture even when considering all known facts. Moreover, the high complexity of these models often makes them not amenable for analysis. Therefore, this motivates the construction of heuristic models that are less complex, but still capture the essential features of a neurological event.

Heuristic models are constructed by including assumed relevant facts to describe a neurological phenomenon of interest via a dynamical system. Since there is no universal agreement over which facts are relevant to a particular phenomenon, the literature for such models is large (see [39] for a review) and more work needs to be done to integrate theoretical neuroscience with experimental work to assess the realism of these models. Neural modelling of this type is therefore more of an art. As our measurement of choice is the electroencephalogram (EEG), which best measures the behaviour of a population of neurons, we consider models that capture the temporal dynamics of neural populations. These are coined in [39] as neural mass models and are governed by ordinary differential equations.

Our motivation lies in anticipating the occurrence of epileptic seizures from the EEG, which is a phenomenon captured by a member of a class of neural mass models that share the same mathematical structure. Other members of this class describe neurological events that include but are not limited to the generation of alpha rhythms and the generation of evoked potentials due to a visual input in the cerebral cortex. We describe them in greater detail in Section 2.3.

2.3 A class of neural mass models

In this section, we present the class of neural mass models we consider, which includes the following models that share the same mathematical structure in their dynamics: (i) The model by Wendling et. al. in [155] that captures epileptic activity in the hippocampus, (ii) The model by Jansen and Rit in [74] on the generation of evoked potentials due to visual stimulation, and (iii) The model by Stam et. al. in [138] for the generation of alpha rhythms. These models have their origins in cortical column models by Wilson and Cowan [156], Freeman [50] and Lopes da Silva et. al. [99; 100]. In the proceeding sections, the class of neural mass models are presented with decreasing level of complexity. The functional connections between the neural populations for each of the models and their corresponding more detailed block diagrams are shown in Figure 2.2-2.7.

2.3.1 Neural mass model by Wendling et. al.

Wendling et. al. built upon the Jansen and Rit model described in the Section 2.3.2. Four neural populations (with one population being a subset of another) are included in this model as shown in Figure 2.2.

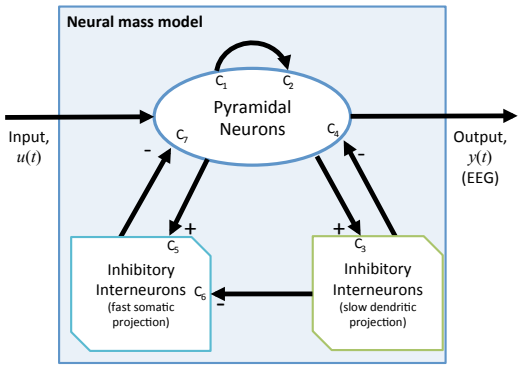


Figure 2.2: Functional relationship between neural populations for the model by Wendling et. al..

The populations are the pyramidal neurons, the excitatory population (included in the pyramidal neurons), the slow and fast inhibitory populations. The fast somatic projection of the inhibitory interneurons is introduced in this model because it is hypothesised to play a role in the fast oscillatory pattern seen in the EEG at the onset of an epileptic seizure. Wendling et. al. identified three parameters, namely the synaptic gains of the excitatory, slow inhibitory and fast inhibitory interneurons to result in the

model producing EEG patterns that are known to be related to neurological events, from normal background activity to epileptic seizures. This provides great motivation for estimating these parameters.

Figure 2.3 describes the interaction between populations of neurons in greater detail, which consists of postsynaptic membrane potential (PSP) kernels h_e, h_i and h_g, sigmoid functions $S : \mathbb{R} \to \mathbb{R}$ and connectivity constants C_1 to C_7.

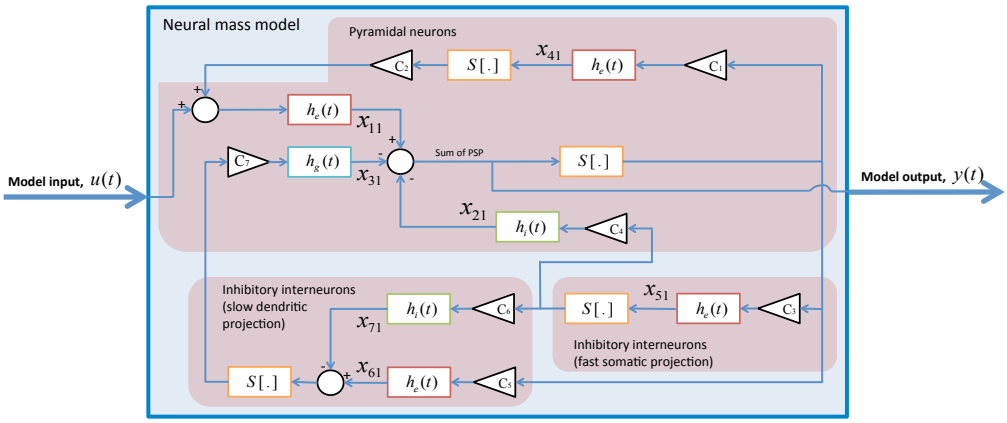


Figure 2.3: Detailed block diagram of the Wendling et al. model. Reproduced from Figure 4 in [155].

The firing rate of the afferent population is converted into an excitatory, slow or fast inhibitory postsynaptic membrane potential via the following kernels, for $t \geq 0$:

- The excitatory population:

$$h_e(t) = \theta_A a t \exp(-at). \tag{2.1}$$

- The slow inhibitory population:

$$h_i(t) = \theta_B b t \exp(-bt). \tag{2.2}$$

- The fast inhibitory population:

$$h_g(t) = \theta_G g t \exp(-gt). \tag{2.3}$$

Parameters θ_A, θ_B and θ_G in (2.1)-(2.3) correspond to the synaptic gains of the excitatory, slow and fast inhibitory populations respectively. These parameters characterise

the observed pattern in the EEG. For example, the values of θ_A, θ_B and θ_G that distinguish between seizure and non-seizure activities have been identified in [155].

Internal variables $x_{11}, \ldots, x_{71}$ are introduced as shown in Figure 2.3. They describe the membrane potential contribution from one population to another. For example, referring to Figure 2.3, the mean membrane potential of the pyramidal neurons is $x_{11} - x_{21} - x_{31}$, which reflects the membrane potential contribution from the excitatory, slow and fast inhibitory populations, respectively. The mean membrane potential of a population is converted into the average firing rate of all the neurons in that population using a sigmoid function S:

$$S(z) = \frac{\alpha_2}{1 + \exp\big(-r_2(z - V_2)\big)}, \qquad \text{for } z \in \mathbb{R}, \tag{2.4}$$

where α_2 is the maximum firing rate of the population, r_2 is the slope of the sigmoid and V_2 is the threshold of the population's mean membrane potential.

The neural populations are connected with connectivity strengths C_1 to C_7, which represents the average number of synaptic contacts between the neural populations concerned.

2.3.2 Neural mass model by Jansen and Rit

The interactions between the pyramidal neurons, excitatory and inhibitory populations (Figure 2.4) are described in this model to investigate the generation of evoked potentials in the cerebral cortex. A detailed block diagram is provided in Figure 2.5. As this model was extended upon by Wendling et. al. (whose model is presented in Section 2.3.1), each component of the block diagram: the PSP kernels h_e and h_i and sigmoidal function S are as described in Section 2.3.1.

2.3.3 Neural mass model by Stam et. al.

The model by Stam et. al. includes an excitatory and inhibitory population (Figure 2.6) to replicate the alpha rhythms in the EEG. This event is related to the human subject being in a relaxed state with the eyes closed. Hence, the estimation of the unmeasured postsynaptic potential (PSP) of neural populations may better our understanding of the visual pathway while in an idle state.

Figure 2.7 shows a detailed block diagram of the model. This model differs from the models by Wendling et. al. and Jansen and Rit in the sense that the firing rate of a population is converted to a postsynaptic potential via different kernels from (2.1)

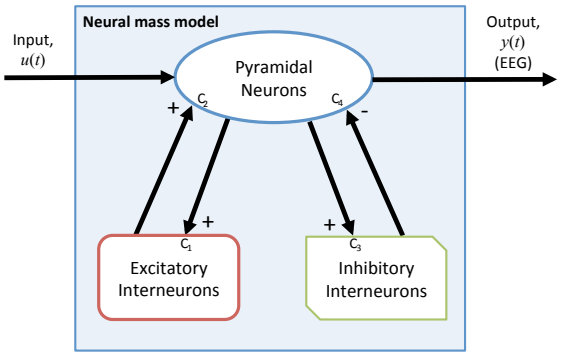


Figure 2.4: Functional relationship between neural populations for the model by Jansen and Rit.

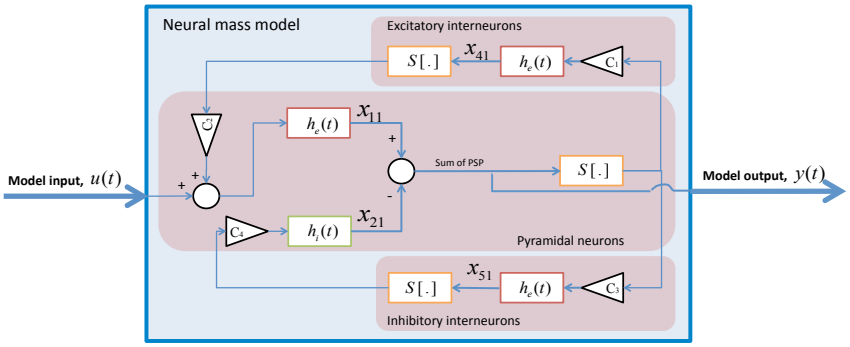


Figure 2.5: Detailed block diagram of the Jansen and Rit model.

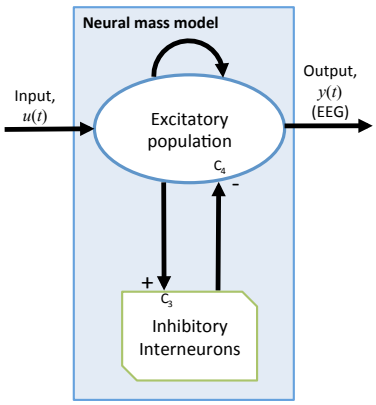


Figure 2.6: Functional relationship between neural populations for the model by Stam et. al..

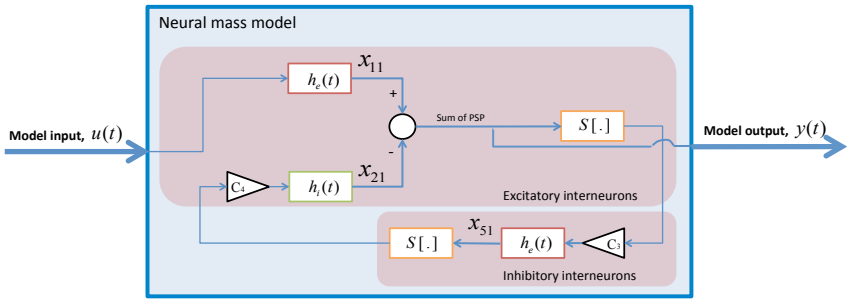


Figure 2.7: Detailed block diagram of the Stam et. al. model.

and (2.2), for $t \geq 0$:

- Excitatory population:

$$h_e(t) = \theta_A[\exp(-a_1 t) - \exp(-a_2 t)]. \tag{2.5}$$

- Inhibitory population:

$$h_i(t) = \theta_B[\exp(-b_1 t) - \exp(-b_2 t)]. \tag{2.6}$$

Also, the sigmoid function that converts the postsynaptic potential to the firing rate of the population differs from (2.4) for the models by Wendling et al. and Jansen and Rit, as follows:

$$S_1(z) = \begin{cases} \alpha_1 \exp\big(r_1(z - V_1)\big) & z \leq V_1, \\ \alpha_1\Big(2 - \exp\big(-r_1(z - V_1)\big)\Big) & z > V_1, \end{cases} \tag{2.7}$$

where α_1 is the maximum firing rate of the population, r_1 is the slope of the sigmoid and V_1 is the threshold of the population's mean membrane potential.

2.4 Neural mass models in state space form

Our observer design is most conveniently carried out using the state space form of the neural mass models. However, some of the neural mass models of interest presented in Section 2.3 were in block diagram form (e.g. Figure 2.3, 2.5 and 2.7) in the given references [74; 138; 155]. In [74] and [155], state space forms were provided but not in the convenient state coordinates where the techniques we use for proving convergence

of estimates can be applied. Therefore, we illustrate how this can be done for the model by Wendling et al., whose detailed block diagram can be found in Figure 2.3. The other models considered are special cases of the model by Wendling et al. and hence, can easily be obtained from the derivation below.

We will show that all neural mass models from Section 2.3 can be written in the following state space form:

$$\dot{x} = Ax + G(p^\star)\gamma(Hx) + \sigma(u, Cx, p^\star), \tag{2.8}$$

and the output of the model is:

$$y = Cx, \tag{2.9}$$

where the state vector is $x \in \mathbb{R}^{n_x}$, input is $u \in \mathbb{R}$, output/EEG measurement is $y \in \mathbb{R}$, parameter vector is $p^\star \in \mathbb{R}^{n_p}$, $G : \mathbb{R}^{n_p} \to \mathbb{R}^{n_x \times m}$, nonlinearity $\gamma = (\gamma_1, \ldots, \gamma_m)$ with $\gamma_i : \mathbb{R} \to \mathbb{R}$ for $i \in \{1, \ldots, m\}$ and nonlinearity $\sigma = (\sigma_1, \ldots, \sigma_n)$ with $\sigma_i : \mathbb{R} \times \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}$ for $i \in \{1, \ldots, n\}$. The number of states n_x, number of parameters n_p and number of scalar nonlinear functions m differs for each model. These are defined in Sections 2.4.2-2.4.4.

2.4.1 Physiological interpretation

Physiologically, the first term in (2.8) implements the postsynaptic potential (PSP) kernels from (2.1), (2.2) and (2.3). This is effectively a convolution of the pre-synaptic firing rates arriving from other populations with the appropriate PSP response functions. These firing rates are modelled in the second and third term in (2.8) that incorporates the sigmoid firing rate function of the depolarisation of contributing populations. The second term, $G(p)\gamma(Hx)$, reflects the influence of all states except the membrane potential of pyramidal population Cx. While the third term, $\sigma(u, Cx, p)$, reflects the influence of the mean membrane potential of the pyramidal cells Cx and the exogenous input u. Neurobiologically, $G(p)\gamma(Hx) + \sigma(u, Cx, p)$ correspond to the effects of intrinsic and extrinsic connections. In other words, when coupling different neural mass models, one has to consider the (intrinsic) influence of populations within a neural mass model and (extrinsic) contributions from other neural mass models. The extrinsic contributions are usually mediated through pyramidal cell populations.

In the following sections, we present the state space form for each model for ease of observer design. Detailed derivations are first shown for the model by Wendling et

al., as it is the most complex model and then the subtle differences in derivations are described for the other models.

2.4.2 State space form for the model by Wendling et al.

We write the Wendling et al. model in state space form by introducing the state variables x_{i1} for $i \in \{1, \ldots, 7\}$ as the membrane potential contribution from one population to another and x_{i2} for $i \in \{1, \ldots, 7\}$ as its derivative. The states x_{i1} are introduced at the outputs of all the impulse responses h_e, h_i and h_g blocks as shown in Figure 2.3. Recalling that the Laplace transform of the impulse responses h_e, h_i and h_g (as described by (2.1), (2.2) and (2.3)) are second-order transfer functions, by performing the inverse Laplace transform, each transfer function is represented by a second-order ordinary differential equation (ODE). We show this transformation for h_e from (2.1) as an example. Let the input to the h_e block be $\bar{u}$ and output be $\bar{y}$. We denote the Laplace transform of signal v as $\mathcal{L}(v)$. Hence, the Laplace transform of h_e with zero initial conditions is:

$$\mathcal{L}(h_e(t)) = \mathcal{L}(\theta_A a t \exp(-at)) = \frac{\theta_A a}{(s+a)^2}. \tag{2.10}$$

Recalling that $\mathcal{L}(h_e) = \frac{\mathcal{L}(\bar{y})}{\mathcal{L}(\bar{u})}$, we obtain:

$$\mathcal{L}(\bar{y})s^2 + 2a\mathcal{L}(\bar{y})s + a^2\mathcal{L}(\bar{y}) = \theta_A a \mathcal{L}(\bar{u}). \tag{2.11}$$

By taking the inverse Laplace transform, we obtain a second-order ODE as follows:

$$\ddot{\bar{y}} + 2a\dot{\bar{y}} + a^2\bar{y} = \theta_A a \bar{u}. \tag{2.12}$$

The x_{i2} states are defined as $x_{i2} = \dot{x}_{i1}$ for $i \in \{1, \ldots, 7\}$ to rewrite the second-order ODEs as two first-order ODEs for each impulse response block.

We illustrate this for the $h_e(t)$ block in the fast inhibitory population, then the output of that block is $\bar{y} = x_{51}$ and the input is $\bar{u} = C_3 S(x_{11} - x_{21} - x_{31})$. Taking $x_{52} = \dot{x}_{51}$, (2.12) can be written as two first-order ODE as follows:

$$\begin{aligned} \dot{x}_{51} &= x_{52} \\ \dot{x}_{52} &= -2a x_{52} - a^2 x_{51} + \theta_A a C_3 S(x_{11} - x_{21} - x_{31}). \end{aligned}$$

Hence, each impulse response h_e, h_i and h_g will each introduce a first-order ODE

in the following general state space form by taking $x_i = (x_{i1}, x_{i2})$ for $i \in \{1, \ldots, 7\}$:

$$\dot{x}_i \quad = \quad A_i x_i + (0, \vartheta_i S(\mu_i) + \varphi_i)\,, \tag{2.13}$$

where μ_i is the input to the respective sigmoid functions, $A_i = \begin{bmatrix} 0 & 1 \\ -k_{i1}k_{i2} & -(k_{i1}+k_{i2}) \end{bmatrix}$, for $i = \{1, \ldots, 7\}$ with $k_{11} = k_{41} = k_{51} = k_{61} = a$, $k_{12} = k_{42} = k_{52} = k_{62} = a$, $k_{21} = k_{71} = b$, $k_{22} = k_{72} = b$, $k_{31} = g$ and $k_{32} = g$. ϑ_i and φ_i are defined as such $\vartheta_1 = \theta_A a C_2$, $\vartheta_2 = \theta_B b C_4$, $\vartheta_3 = \theta_G g C_7$, $\vartheta_7 = \theta_B b C_6$, $\vartheta_4 = \vartheta_5 = \vartheta_6 = 0$ and $\varphi_1 = \theta_A a u$, $\varphi_2 = \varphi_3 = \varphi_7 = 0$, $\varphi_4 = \theta_A a C_1 S(y)$, $\varphi_5 = \theta_A a C_3 S(y)$, $\varphi_6 = \theta_A a C_5 S(y)$. Constants a, b and g are strictly positive. S is a sigmoid function described by (2.4). All constants discussed in this section are summarised in B.12.

The subsystems defined in (2.13) are put together to be written compactly in state space form (2.8)-(2.9) for ease of observer design.

We take the state vector in (2.8) and (2.9) to be $x = (x_1, \ldots, x_7)$ where x_i for $i = \{1, \ldots, 7\}$ satisfy (2.13). The states x_1, x_2 and x_3 capture the membrane potential contribution and its derivative of the excitatory, slow and fast inhibitory populations to the pyramidal neurons, respectively. The states x_4, x_5 and x_6 capture the membrane potential contribution and its derivative of the pyramidal neurons to the excitatory, slow and fast inhibitory populations, respectively. The output is $y = x_{11} - x_{21} - x_{31}$. The specific matrices in (2.8) and (2.9) are denoted as:

- The parameter vector is $p^\star = (\theta_A, \theta_B, \theta_G)$,
- The matrix $A = \mathrm{diag}(A_1, \ldots, A_7)$,
- $\gamma = (S, S, S)$, where S is defined in (2.4),
- $\sigma = (0, \theta_A a u, 0, 0, 0, 0, 0, \theta_A a C_1 S(y),\ 0, \theta_A a C_3 S(y), 0, \theta_A a C_5 S(y), 0, 0)$, where S is described by (2.4),
- $C = [\ 1 \quad 0 \quad -1 \quad 0 \quad -1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0\ ]$,

- $G = \begin{bmatrix} 0 & 0 & 0 \\ \theta_A a C_2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \theta_B b C_4 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \theta_G g C_7 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \theta_B b C_6 & 0 \end{bmatrix},$

- $H = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \end{bmatrix}.$

2.4.3 State space form for the model by Jansen and Rit

We write the model in state space form by taking the state vector in (2.8) to be $x = (x_1, x_2, x_4, x_5)$, where x_i for $i = \{1, 2, 4, 5\}$ satisfy (2.13). States x_1 and x_2 are the membrane potential contribution and its derivative of the excitatory and inhibitory populations to the pyramidal neurons, respectively. States x_4 and x_5 capture the membrane potential contribution and its derivative of the pyramidal neurons to the excitatory and inhibitory populations, respectively. The output is $y = x_{11} - x_{21}$. The specific matrices in (2.8) and (2.9) are denoted as:

- The parameter vector is $p^\star = (\theta_A, \theta_B, C_1, C_2, C_3, C_4)$,
- $A = \text{diag}(A_1, A_2, A_4, A_5)$,
- $\gamma = (S, S)$ where S is defined in (2.4),
- $\sigma = (0, \theta_A a u, 0, 0, 0, \theta_A a C_1 S(y), 0,\ \ \theta_A a C_3 S(y))$,
- $C = [\ 1 \quad 0 \quad -1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0\ ]$,

- $G = \begin{bmatrix} 0 & 0 \\ \theta_A a C_2 & 0 \\ 0 & 0 \\ 0 & \theta_B b C_4 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$,
- $H = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$.

2.4.4 State space form for the model by Stam et al.

The model is written in state space form by taking the state vector in (2.8) as $x = (x_1, x_2, x_5)$ where x_i for $i = \{1, 2, 5\}$ satisfy (2.13). State $x_1 = (x_{11}, x_{12})$ represents the mean membrane potential of the excitatory population's activity to itself and its derivative, respectively. States $x_2 = (x_{21}, x_{22})$ and $x_5 = (x_{51}, x_{52})$ represent the mean membrane potential and its derivative of the inhibitory population to the excitatory population and vice versa, respectively. The output is $y = x_{11} - x_{21}$.

As mentioned in Section 2.3.3, the PSP kernels h_e and h_i differ from the ones in the models by Wendling and Jansen et. al.. Nevertheless, rewriting (2.5)-(2.7) into state space form does not differ from the derivation presented in Section 2.4. By following the procedure described in Section 2.4.2, i.e. by taking Laplace transformations of (2.5) and (2.6) and taking the inverse Laplace transform, the kernels can be written as second-order ODEs. They can then be rewritten as two first-order ODEs by introducing extra state variables, x_{i2} for $i \in \{1, 2, 5\}$, in a similar fashion as in Section 2.4.2. The specific matrices in (2.8) are denoted as:

- The parameter vector is $p^\star = (C_3, C_4)$,
- $A = \text{diag}(A_1, A_2, A_5)$,
- $\gamma = S_1$ where S is from (2.7),
- $\sigma = (0, \theta_A(a_2 - a_1)u, 0, 0, 0, \theta_A(a_2 - a_1)C_3 S_1(y))$,
- $C = \begin{bmatrix} 1 & 0 & -1 & 0 & 0 & 0 \end{bmatrix}$,
- $G = (0, 0, 0, \theta_B(b_2 - b_1)C_4, 0, 0)$,

- $H = [\ 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0\ ]$.

2.5 Summary

Three examples from a class of neural mass models are presented in this chapter. They are the models by Wendling et. al., Jansen and Rit as well as Stam et. al. respectively. A distinct feature of the neural mass models considered in this chapter is that they are parameterised models, where the parameters dictate the type of output that is known to represent certain brain activities, such as alpha rhythms in the visual system and epileptic activity. It is known that the brain transitions into different activities in time, such as from non-seizure to seizure conditions. Hence, a more accurate representation would involve assigning dynamics to parameters, where the dynamics are likely to involve the states. One such model was done by Suffczynski et. al. [141] to explain the generation of generalised absence seizures, a form of whole-brain epilepsy. We did not consider such models in this thesis.

The models discussed in this chapter are cortical column models. Cortical columns are repeated all over cortical brain surfaces, such as in the cerebral cortex or the hippocampus. Therefore, cortical column models can be used to describe an arbitrary volume of the brain, a feature useful for describing localised neurological events. Often, neurological events also have spatial dynamics, such as the spread of seizures from a foci. These events have dynamics that are governed by partial differential equations and are known as neural field models [39]. These models are not in the scope of this thesis. Another way of capturing the spatial activity of a neurological event is to interconnect the single cortical column models to characterise the phenomenon in a cortical area, such as the primary visual cortex or even across a larger brain volume. We take this approach in Chapter 5 where we design an adaptive observer that can estimate the mean membrane potential (states) as well as the synaptic gains (parameters) of the neural populations for each of the interconnected single cortical column models.

In Section 2.4 of this chapter, we write each of these example models in a form that is most amenable for the design of state observers in Part I as well as state and parameter observers in Part II of this thesis.

Part I: State estimation

In this part, the estimation of states under the assumption that the parameter is known is performed using the setup illustrated in Figure 2.8. The dynamical system that computes an estimate of the state x of the system Σ is called a **state observer**.

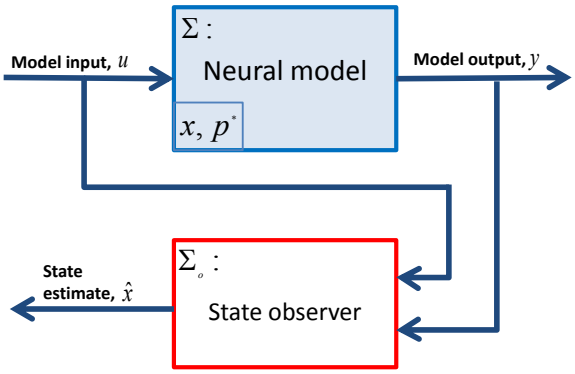


Figure 2.8: State observer setup

As mentioned in the introduction, two important issues need to be addressed in state observer design, which are also applicable to state and parameter observer design. Firstly, do the state estimates $\hat{x}$ (similarly, parameter estimates $\hat{p}$) converge to their true values? Secondly, how do the observers perform in the presence of modelling error, measurement and input noise?

In response to the first set of questions, recall that the neural model (or system) as stated in (1.1) in Chapter 1 takes the following form

$$\begin{aligned} \Sigma : \qquad \dot{x} &= f(x, u, p^*) \\ y &= h(x, p^*), \end{aligned} \tag{2.14}$$

where the state is $x \in \mathbb{R}^{n_x}$, the input is $u \in \mathbb{R}^{n_u}$, the parameter is $p^* \in \Theta \subset \mathbb{R}^{n_p}$, measurement/output is $y \in \mathbb{R}^{n_y}$ and n_x, n_u, n_p and n_y are positive integers.

A state observer for the model (2.14) takes the following form

$$\begin{aligned} \dot{\hat{x}} &= \hat{f}(\hat{x}, u, p^*, y) \\ \hat{y} &= h(\hat{x}, p^*), \end{aligned} \tag{2.15}$$

where the state estimate is $\hat{x} \in \mathbb{R}^{n_x}$, the input is $u \in \mathbb{R}^{n_u}$ and the output estimate is $\hat{y} \in \mathbb{R}^{n_y}$. Our task is to design $\hat{f}$ such that the state estimate converges to its true value.

To achieve this objective, we consider the state estimation error system (similarly, the parameter estimation error system. However, we concentrate only on state estimation in this part.). Denoting the state error as $\tilde{x} := \hat{x} - x$, we obtain the following state error system from (2.14) and (2.15):

$$\dot{\tilde{x}} = \hat{f}(\tilde{x} + x, u, p^*, y) - f(x, u, p^*). \tag{2.16}$$

We define two types of observers based on the convergence properties of their estimates via the stability of the equilibrium point of the state estimation error system (2.16).

Definition 1. *If the origin of the state error system* (2.16) $\tilde{x} = 0$ *is:*

- *exponentially stable*[1]*, the observer* (2.15) *is termed* exponentially convergent. *If the conditions hold for all* $\tilde{x}(0) \in \mathbb{R}^{n_x}$*, the observer* (2.15) *is called* globally *exponentially convergent.*

- *asymptotically stable*[1]*, the observer* (2.15) *is termed* asymptotically convergent *and if the conditions stated are satisfied for all* $\tilde{x}(0) \in \mathbb{R}^{n_x}$*, the observer* (2.15) *is* globally *asymptotically convergent.*

Given general nonlinear systems, the design of convergent observers are reliant on the specific structures of the systems considered. The simplest case being error systems that are linear. For example, consider the nonlinear system $\dot{x} = Ax + \gamma(y)$, $y = Cx$, where γ is a nonlinear function. An observer $\dot{\hat{x}} = A\hat{x} + \gamma(y) + L(C\hat{x} - y)$ can be designed and the resulting error system $\dot{\tilde{x}} = (A + LC)\tilde{x}$ is linear. Given that the pair (A, C) is detectable, the observer matrix L can be chosen such that the matrix $A + LC$ is Hurwitz, which ensures that the designed observer is exponentially convergent. In cases where a general nonlinear system is considered and suppose some coordinate transformation exists such that the resultant state error is linear, then a convergent observer can also be achieved [72, Chapter 4.9], [90; 91].

Another well-known nonlinear observer is the high-gain type, which extends the aforementioned idea of linear error systems by using a high gain to compensate for the nonlinear part such that the resulting error system is 'almost' linear [13; 29; 54; 123].

[1] see Definition 2 in Appendix A

As with all other observer designs for nonlinear systems, the high gain observer is only applicable to a class of nonlinear systems with special structures. Consider for example the nonlinear system [23, Chapter 2.2.2]:

$$\begin{aligned} \dot{x} &= Ax + \gamma(u, x), \\ y &= Cx, \end{aligned} \tag{2.17}$$

where γ is globally Lipschitz, $C = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}$, $A = \begin{bmatrix} 0 & 1 & & 0 \\ \vdots & & \ddots & \\ & & & 1 \\ 0 & \dots & & 0 \end{bmatrix}$ and $\gamma(u, x) = \begin{bmatrix} \gamma_1(x_1, u) \\ \gamma_2(x_1, x_2, u) \\ \vdots \\ \gamma_n(x, u) \end{bmatrix}$ (triangular form).

The high gain observer takes the form $\dot{\hat{x}} = A\hat{x} + \gamma(u, \hat{x}) + \Delta K(C\hat{x} - y)$, where $\Delta = \text{diag}(d, d^2, \dots, d^n)$, with the parameter $d > 0$ to be chosen and the observer matrix K to be designed. By considering a scaled state error, $\bar{x} := \Delta^{-1}\tilde{x}$, the resulting scaled error system is $\dot{\bar{x}} = d(A + KC)\bar{x} + \Delta^{-1}(\gamma(u, \hat{x}) - \gamma(u, x))$. The parameter $d > 0$ can then be chosen sufficiently large such that the origin of the scaled error system is exponentially stable [23, Chapter 2.2.2]. Such ideas have also been employed in the design of *adaptive* observers [25; 26; 46], of which we will be adapting to suit the class of neural mass models we consider in Chapter 5.

Two notable drawbacks of the high gain design are that it relies heavily on the global Lipschitzness and the triangular structure of the nonlinearity γ. The circle criterion observer introduced by Arcak and Kokotović [16] removes these restrictions. Consider now a class of nonlinear systems $\dot{x} = Ax + G\gamma(Hx) + \sigma(y, u)$ and the proposed circle criterion observer takes the form $\dot{\hat{x}} = A\hat{x} + G\gamma(H\hat{x} + K(C\hat{x} - y)) + L(C\hat{x} - y) + \sigma(y, u)$. The error system is viewed as Lure's system (Figure 2.9) and two restrictions are placed on the error system in order to satisfy the well-known multi-variable circle criterion [86, Theorem 7.1]. The two restrictions are that the linear part of the error system $H(s)$ is strictly positive real (SPR) and the nonlinearity γ is monotonically increasing. The satisfaction of the SPR condition is implied by the feasibility of a linear matrix inequality (LMI). Several relaxations on the system have been performed to obtain a less conservative LMI, including introducing a multiplier [45] or by considering a globally Lipschitz nonlinearity γ [17; 158]. For the class of neural mass models considered, the

original circle criterion observer introduced in [16] resulted in an LMI that is infeasible. So, in Chapter 4, we have combined these methods to obtain a state observer that is applicable to our class of models.

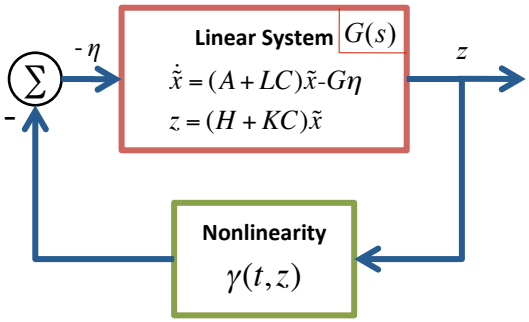


Figure 2.9: Circle criterion observer: error system

The second pertinent issue that arises in the estimation problem is the robustness property of the observer towards modelling error, as well as noise in input and measurements. In the case of the state observer for a family of parameterised systems such as the class of neural mass models we consider, where the parameters are assumed to be slowly-varying such that it can be considered constant for the observation period, modelling error takes the form of the error between the assumed parameter and the real parameter. The robustness of the observer towards perturbations in parameter is the basis that motivates the next step, that is to jointly estimate the parameters as well.

For the practical implementation of observers in the context of estimating neural activity using the EEG, robustness towards measurement noise is a highly desirable property. The measured EEG is recorded using scalp or intracranial electrodes, which usually contains artefacts from the physical movement of the sensor or interference from other electrical sources such as the mains. Hence, in the design of our observers, we will take these issues into account by either showing *a posteriori* robustness or by incorporating tuneable design parameters to improve robustness. An important concept that is applied to show or improve robustness is the notion of *input-to-state* stability (ISS), first introduced by Sontag [131] (see [133] for overview of recent results). In fact, an interesting application of this concept is also employed in the design of a nonlinear estimator for the class of neural mass models in Chapter 2.

Overview of Part I

The next two chapters focus on estimating the states of the class of systems presented in Chapter 2 using a setup as shown in Figure 2.8. Recall from (2.8) that we consider the following class of systems:

$$\begin{aligned} \dot{x} &= Ax + G(p^*)\gamma(Hx) + \sigma(u, y, p^*) \\ y &= Cx, \end{aligned} \tag{2.18}$$

where the state vector is $x \in \mathbb{R}^{n_x}$, the input is $u \in \mathbb{R}^{n_u}$, the measurement is $y \in \mathbb{R}^{n_y}$, $\gamma = (\gamma_1, \ldots, \gamma_m) : \mathbb{R}^{n_m} \to \mathbb{R}^{n_m}$ and $\sigma = (\sigma_1, \ldots, \sigma_n) : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_p} \to \mathbb{R}^{n_x}$.

Assuming that the parameters p^* are known, the nonlinear state observers designed in this part take the following form:

$$\dot{\hat{x}} = A\hat{x} + G(p^*)\gamma(H\hat{x} + K(C\hat{x} - y)) + L(C\hat{x} - y) + \sigma(u, y, p^*), \tag{2.19}$$

where $\hat{x} \in \mathbb{R}^n$ is the state estimate and observer matrices $K \in \mathbb{R}^m$, $L \in \mathbb{R}^n$ are introduced via linear output injection terms $K(C\hat{x} - y)$ and $L(C\hat{x} - y)$. This structure, in particular the output injection term $K(C\hat{x} - y)$, was first proposed in the design of circle criterion observers in [16]. We adopted this observer structure in our designs proposed in Chapters 3 and 4, where the observer matrices K and L are designed via different methodologies.

We first propose a class of nonlinear observers that are designed specifically for the class of neural mass models presented in Chapter 2. By noting that the observation error systems of this class of neural mass models has a cascade property, we use an interesting application of the concept of ISS in the stability analysis of a cascade system to show that the designed observers are globally exponentially convergent. This class of nonlinear observers is presented in Chapter 3.

Next, we extend the circle criterion observer first designed by Arcak and Kokotović in [16] such that it can be applied to the class of neural mass models we consider. We also allow the user to tune the attenuation factors of measurement noise and input uncertainty with the feasibility of the observer as a tradeoff. We call this a *robust* circle criterion observer. The circle criterion observers proposed in Chapter 4 are applicable to *any* system in this class (2.18) and are not limited to neural mass models.

Chapter 3

A nonlinear observer

We propose a nonlinear observer for the class of neural mass models considered in Chapter 2 by first considering the idealised conditions stated in Assumptions 1-4. The proposed observer is considered to be of the extended Luenberger type due to the introduction of two linear output injection terms in the dynamics of the observer. The novelty lies in the linear output injection term that appears in the nonlinearity, a concept first introduced for circle criterion observers in [16]. The nonlinear observer presented in this chapter differs from circle criterion observers in that the observer matrices are designed differently, via a methodology that relies heavily on the mathematical properties of the neural mass models considered, i.e. a linear part that is stable, nonlinearities that are globally Lipschitz and a state estimation error system that can be decomposed into subsystems with a cascade structure.

The idealised assumptions are later relaxed and we investigate the robustness of the proposed observer against disturbances and uncertainties. Simulation results are provided to illustrate the convergence of the state estimates and the robustness of the designed observer.

Recall from Chapter 2.4 that the neural mass models of interest can be written in the form:

$$\begin{aligned} \dot{x} &= Ax + G(p^\star)\gamma(Hx) + \sigma(u, Cx, p^\star), \\ y &= Cx. \end{aligned} \tag{3.1}$$

We assume the following about (3.1):

Assumption 1. *The synaptic gain of each neuronal population $\theta_A, \theta_B, \theta_G$ and the connectivity strengths C_1, C_2, C_3, C_4 are parameters that are constant and known.*

These parameters are typically slowly-varying during a particular brain activity, such that they can be considered constant over the time period observed. When the brain transitions from one activity to another however, the parameter $p^\star = (C_3, C_4)$ for the model by Stam et al. in [138], parameter $p^\star = (\theta_A, \theta_B, C_1, C_2, C_3, C_4)$ for the model by Jansen and Rit in [74] and parameters $p^\star = (\theta_A, \theta_B, \theta_G)$ for the model by Wendling et al. in [155] will vary. If the goal is solely to estimate the parameters, system identification methods such as least-squares estimation [98] or genetic algorithms [73] may be used. However, it remains a hard problem due to the nonlinearity of these models. In Part II of this thesis, we propose two parameter and state estimation methods. In this chapter, we only consider the case where the brain is in one particular brain activity and that the parameters are known, that is for the case where the parameters have been identified a priori.

Assumption 2. *The input u is measured.*

The input from afferent populations is hard to quantify in practice, hence this assumption is not justified in general. We will later relax this assumption by allowing the input to be uncertain. We show that good estimates can still be obtained provided that the L_∞ norm of the difference between the true and assumed input is small. It is important to note that it is not an easy task to guarantee exponential convergence of the error estimates without assuming that the input is known exactly. Please refer to [61] and references therein.

Assumption 3. *The measured EEG, y is noise-free.*

The measured EEG considered here can be recorded using both scalp or intracranial electrodes (electrocorticography or ECoG). For best results, the observer is best applied to ECoG, which has the highest signal-to-noise ratio such that it can be considered noise-free. We will show later that the estimation error remains bounded despite noisy measurements. We later relax our assumptions and show that the designed observer is robust to measurement noise.

Assumption 4. *The model dynamics are accurate.*

We later allow for model uncertainty in Section 3.2 by introducing an additive disturbance to the dynamics and show that the estimation error remains bounded with bounded disturbance.

In Section 3.1, we first synthesise a nonlinear observer in an ideal setting (Assumptions

1-4). We relax these assumptions in Section 3.2 and show that the nonlinear observer is robust towards uncertainties and disturbances. Finally, in Section 3.3, we provide simulations under different scenarios to illustrate the results developed in Sections 3.1 and 3.2.

3.1 Observer design under the ideal scenario

The structure of the nonlinear observer proposed in this chapter shares the same *form* as the circle criterion observer first proposed by Arcak and Kokotović in [16]. However, the *design* of the observer in this chapter differs due to the particular structure of the neural mass models discussed in Chapter 2. We will revisit the circle criterion observers proposed in the literature later in Chapter 4 and adapt them such that they are applicable to our examples. We propose the following nonlinear observer:

$$\dot{\hat{x}} \quad = \quad A\hat{x} + G(p^\star)\gamma(H\hat{x} + K(C\hat{x} - y)) + L(C\hat{x} - y) + \sigma(u, y, p^\star), \tag{3.2}$$

where $\hat{x} \in \mathbb{R}^n$ is the state estimate and observer matrices $K \in \mathbb{R}^m$, $L \in \mathbb{R}^n$ are introduced via linear output injection terms $K(C\hat{x} - y)$ and $L(C\hat{x} - y)$.

The observer matrices L and K in (3.2) are the choice of the user. They are introduced to provide flexibility over the convergence rate of the estimates. We establish in Theorem 1 that for sufficiently small norms of matrices K and L, we obtain convergence of the estimates to the true states in exponential time, for any initial conditions.

Theorem 1. *Consider the model in general form (3.1) under Assumptions 1 to 4 and observer matrices K and L for (3.2) are chosen such that:*

$$\rho|K||G| + |L| \quad < \quad \nu\frac{1}{|C|}, \tag{3.3}$$

where ρ is the Lipschitz constant of γ, matrices G and C are from the model (3.1) and $\nu > 0$ is constructed in the proof.

Then, the observer (3.2) is a global exponential observer for the model (3.1), i.e. for any $u \in L_\infty$ and denoting the estimation error as $\tilde{x} := x - \hat{x}$,

$$|\tilde{x}(t)| \quad \leq \quad k\exp(-\lambda t)|\tilde{x}(0)| \qquad \forall t \geq 0, \forall \tilde{x}(0) \in \mathbb{R}^n, \tag{3.4}$$

where constants k, $\lambda > 0$.

Proof. The proof is given in Appendix B.1. □

Remark 1. *Note that Theorem 1 applies to all neural mass models considered in Chapter 2, where matrices G and C are different for each model considered. Consequently, the obtained observer matrices K and L as well as ν from (3.3) and k, λ from (3.4) differ for each model.*

The exponential convergence property of the observer is desirable in practice because it means that the estimates are guaranteed to converge to the true states in exponential time. Additionally, the validity of the convergence property for any initial conditions is highly desirable because the initial conditions of the neuronal populations are usually unknown. Hence, the observer can be initialised with arbitrary initial conditions.

Theorem 1 provides us with a condition (3.3) that gives a class of nonlinear observers parameterised by matrices K and L. This condition (3.3) is conservatively obtained in B.1. However, we see that the choice of $L = 0$ and $K = 0$ fulfils condition (3.3) and therefore, admits an estimator for the model (2.8) [36]. The drawback of an estimator lies in that the convergence speed of the state estimation error cannot be controlled by the user. Nevertheless, it remains a useful observer provided that the error converges sufficiently fast. In Section 3.3.1, we see in simulations that the estimator for the model by Wendling et al. [155] provides estimates that converge to the true states in a reasonable timeframe. Non-zero choices of K and L fulfilling condition (3.3) provide tuneability of the estimation error's convergence speed. However, condition (3.3) does not provide a priori information about the convergence rate of the observer. In Section 3.3.1, we show in simulations that some choices of non-zero matrices K and L that satisfy condition (3.3) can lead to faster convergence rate for the state estimation error.

Remark 2. *The observer structure (3.2) shares the same mathematical structure as other nonlinear observers in the literature, that is, the high gain [120] (with $K = 0$) and circle criterion observers [16]. These observers employ special techniques in obtaining observer matrices L and K that are not satisfied by the model we consider (2.8). Therefore, we prove the existence of K and L matrices using a different analysis from that in [16; 120].*

3.2 Robustness analysis

Assumptions 1-4 are strong, but they provide us with a good first step in proving the convergence of state estimates. We can relax these assumptions and characterise their effects on the convergence property of the state estimation error system. Figure 3.1

summarises our relaxed assumptions and we restate them as follows:

Assumption 5 (Relaxation of Assumption 1)**.** *The parameters θ_A, θ_B, θ_G, C_1, C_2, C_3 and C_4 of the model are known with error.*

We characterise these uncertainties in parameters by introducing bounded, time-varying signal $\epsilon_p(t)$ to the parameters of the model, so that the true parameter values in the model become $p^\star + \epsilon_p$.

Assumption 6 (Relaxation of Assumption 2)**.** *The input is uncertain.*

We introduce bounded, time-varying disturbances $\epsilon_u(t)$ to the input available to the observer. Thus, the real input is modelled by $u + \epsilon_u$, whereas the observer is fed by an assumed input u. The introduction of ϵ_u relaxes Assumption 2 by allowing the real input to be unknown to the observer.

Assumption 7 (Relaxation of Assumption 3)**.** *The measured EEG is affected by measurement noise.*

Disturbance $\epsilon_y(t)$ is introduced to characterise measurement noise. The measurement with noise is denoted as $y + \epsilon_y$.

Assumption 8 (Relaxation of Assumption 4)**.** *The neural mass model is inaccurate.*

Additive disturbance $\epsilon_{sys}(t)$ characterises the possible influence from other populations to the neuronal populations included in the model.

We show in Theorem 2 that under Assumptions 5-8, i.e. uncertainties in modelling, parameters, measurement and input, the estimates provided are close to the true states, where the 'closeness' is determined by the L_∞ norm of the uncertainties ϵ_y, ϵ_p, ϵ_u and ϵ_{sys}. The introduction of the uncertainties and disturbances to the ideal setup is illustrated in Figure 3.1.

Therefore, we obtain the following perturbed systems (3.5) and (3.6) from (3.1) and (3.2) as illustrated in Figure 3.1:
From (3.1):

$$\begin{aligned} \dot{x} &= Ax + G(p^\star + \epsilon_p)\gamma(Hx) + \sigma(u, Cx, p^\star + \epsilon_p) + \epsilon_{sys}, \\ y &= Cx. \end{aligned} \tag{3.5}$$

From (3.2):

$$\begin{aligned} \dot{\hat{x}} &= A\hat{x} + G(p^\star)\gamma\big(H\hat{x} + K(C\hat{x} - (y + \epsilon_y))\big) \\ &\quad + L(C\hat{x} - (y + \epsilon_y)) + \sigma(u + \epsilon_u, y + \epsilon_y, p^\star). \end{aligned} \tag{3.6}$$

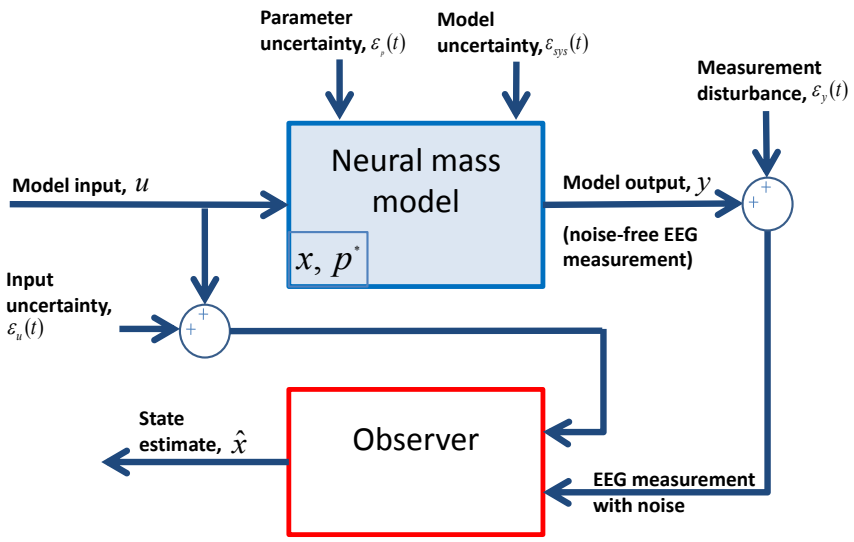


Figure 3.1: Model and observer setup under relaxed assumptions.

Theorem 2. *Consider the perturbed estimation error system (3.5) and (3.6) under Assumptions 5 to 8. The observer (3.6) with matrices K and L are chosen such that (3.3) is satisfied.*

This guarantees that for all $\tilde{x}(0) \in \mathbb{R}^n$, $t \geq 0$ and for all u, ϵ_y, ϵ_p, ϵ_u and $\epsilon_{sys} \in L_\infty$, the error system satisfies the following:

$$\begin{aligned} |\tilde{x}(t)| \quad \leq \quad & \tilde{k}\exp(-\tilde{\lambda}t)|\tilde{x}(0)| + \gamma_y(\|\epsilon_y\|_{[0,t]}) + \gamma_\theta(\|\epsilon_p\|_{[0,t]}) \\ & + \gamma_u(\|\epsilon_u\|_{[0,t]}) + \gamma_{sys}(\|\epsilon_{sys}\|_{[0,t]}), \end{aligned} \tag{3.7}$$

where $\tilde{k}$, $\tilde{\lambda} > 0$ and γ_y, γ_θ, γ_u, $\gamma_{sys} : \mathbb{R} \to \mathbb{R}$ are continuous positive increasing functions that are zero at the origin.

Proof. The proof for Theorem 2 is provided in Appendix B.2. □

We see that with no uncertainties and disturbance, i.e. $\epsilon_u = 0$, $\epsilon_p = 0$ and $\epsilon_y = 0$, we recover the results of Theorem 1. Loosely speaking, Theorem 2 states that small uncertainties and disturbance implies small estimation error.

Remark 3. *By setting $\epsilon_u = -u$, we have that the input to the observer (3.2) is 0, that is we consider the proposed observer (3.2) as an unknown-input observer. The estimation error of the proposed observer (3.2) converges with some error that depends upon the L_∞ norm of ϵ_u. This result is advantageous because the mean firing rate of the*

afferent neuronal populations to a particular brain region is hard to measure in reality. We will further demonstrate this observation in Section 3.3.3 via simulations.

3.3 Simulation

We illustrate the performance of our observers using simulations. First, we perform simulations to show the convergence of the estimates under the ideal conditions stated in Assumptions 1-4 and test the influence of the observation gains on the speed of convergence of the estimation error. Next, we show the robustness of our observers against uncertainties in modelling, parameters, input and measurement in Section 3.3.2. We also show that our observers can provide estimates that are close to to the true states when the input to the observer is unknown in Section 3.3.3. Note that estimating the hidden states over multiple populations within our neural mass model is a non-trivial endeavour, given just a single channel of data. The simulation results provided are for the model of the hippocampus by Wendling et al. in [155]. Similar results can be obtained for other models introduced in Chapter 2.

We select synaptic gains $\theta_A = 5$, $\theta_B = 25$ and $\theta_G = 10$ that correspond to seizure activity as identified in [155]. The initial condition of the model (2.8) is set to $x(0) = [6, 0.5, 6, 0.5, 6, 0.5, 6, 0.5, 6, 0.5, 6, 0.5, 6, 0.5]^T$. We initialise the observer (3.2) to $\hat{x}(0) = 0$. This choice is arbitrary as we have shown that the convergence of the estimation error is valid for all initial conditions in Theorems 1 and 2.

In this section, we denote the observer matrices as $K = k \times I_k$ and $L = l \times I_l$ where $I_k = [1, 1, 1]^T$ and $I_l = [1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1]^T$. We call k and l the observer gains. For each of the scenarios in the following sections, the performance of two observers are evaluated:

1. The estimator $l = 0$, $k = 0$.
2. The observer with observer gains $k = 0.1$ and $l = -0.2$.

All other constants used in simulation are as specified in Appendix B.12.

3.3.1 Simulations under Assumptions 1 to 4: ideal scenario

We consider the ideal case where the parameters are constant and known, input u and measurement y are known and unperturbed as well as no modelling errors. Both model (2.8) and observer (3.2) are supplied with the same Gaussian noise input with mean 90mV and variance 30mV used in [155].

As stated in Theorem 1 and illustrated in Figures 3.2 and 3.3, the state estimation error $e := x - \hat{x}$ converges to 0 asymptotically in exponential time. We observe in simulations that for both observers, the convergence rate is in general faster than the duration of a specific brain activity (see [155]): by $t = 0.3s$, all state estimation errors have converged to 0. As seen in Figure 3.2, the plot of the error norm shows a large overshoot initially (approximately 1200). However, the plots on Figures 3.3 show that the estimates of the membrane potential contributions from one population to another $\hat{x}_{i1}$ are reasonably close to the true values x_{i1}. The large initial error is due to the estimation error in the other states, i.e. the error of the time derivative of the membrane potential contribution $|x_{i2} - \hat{x}_{i2}|$, which are not the physical quantities of concern in practice.

As discussed in Section 3.1, the choice of non-zero observer gains provides flexibility over the convergence rate of the estimation error. Figure 3.2 shows that for a selection of observer gains k and l, the norm of the estimation error converges faster than with the case of $k = l = 0$ (estimator).

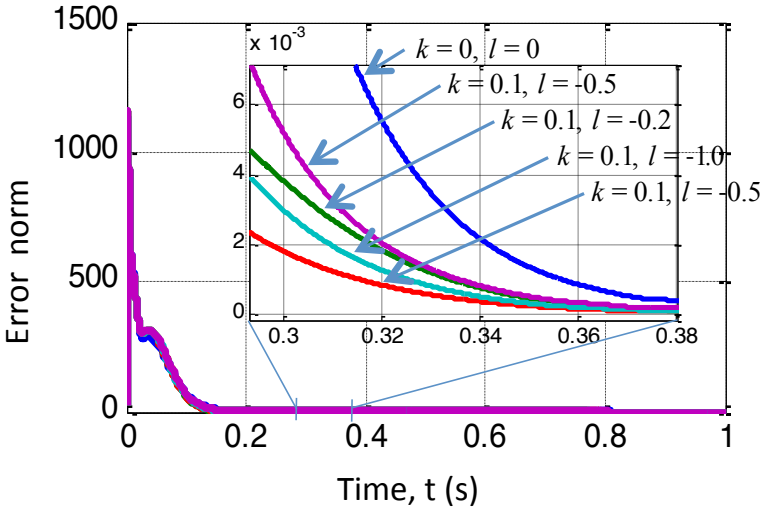


Figure 3.2: Error norm $|\tilde{x}(t)|$ for a selection of k and l.

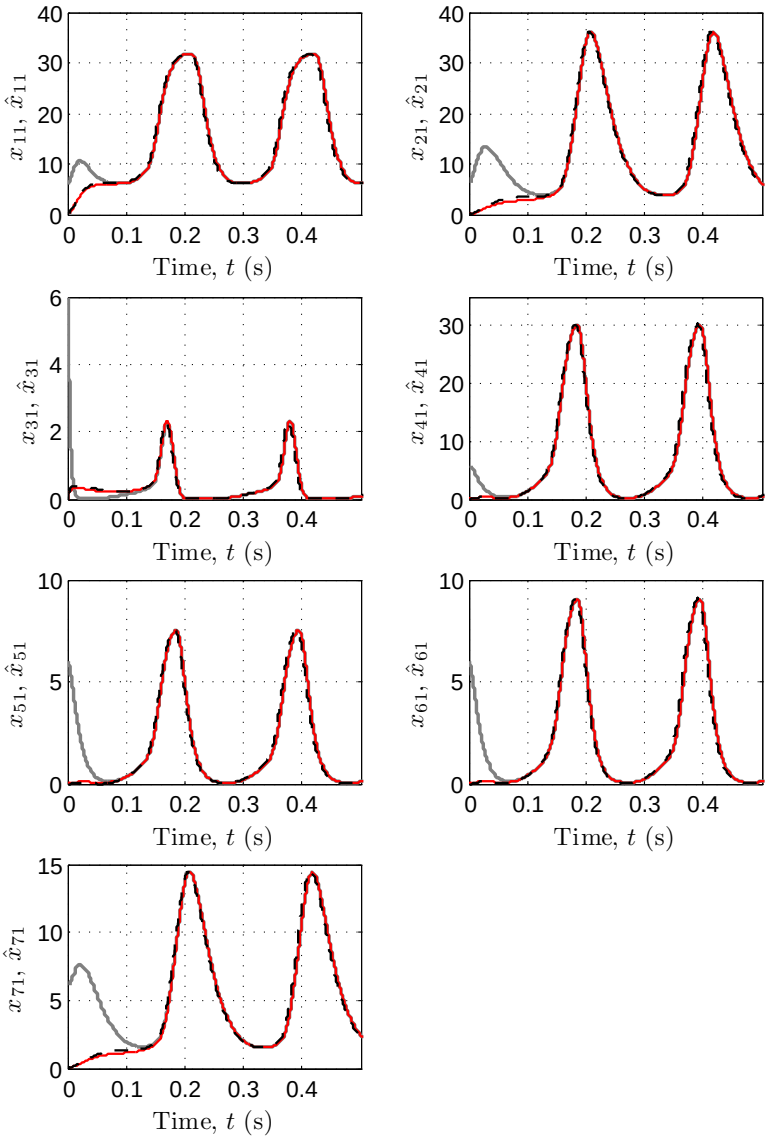


Figure 3.3: Membrane potential x_{i1} (grey solid line) and estimated membrane potential contribution $\hat{x}_{i1}$ (red solid line: $l = k = 0$, black dashed line: $k = 0.1$, $l = -0.5$) under the ideal scenario (Assumptions 1 to 4).

3.3.2 Simulations under Assumptions 5 to 8: uncertainty in parameters, input and measurement as well as additive disturbance.

Next, we test the performance of the observers under uncertainties in the parameters θ_A, θ_B, and θ_G, disturbances in the input and output of the observer, as well as additive disturbance.

We simulate the perturbed systems (3.5) and (3.6) with the perturbations: constant parameter uncertainty $\epsilon_p = (0.5, 2.5, 1)$, Gaussian input uncertainty $\epsilon_u \sim \mathcal{N}(0, 0.3^2)$, Gaussian measurement noise $\epsilon_y \sim \mathcal{N}(0, 0.1^2)$ and Gaussian model uncertainty $\epsilon_{sys} \sim \mathcal{N}(0, 1^2\mathbb{I})$. The performance of both observers are similar. Figures 3.4 and 3.5 show that the estimation error converges to a neighbourhood of the origin. This observation agrees with Theorem 2.

Figure 3.4 shows that the estimates for the states of interest x_{i1} do converge reasonably close to the true states. It can be seen in Figure 3.5 that the relative error of the states is quite low (under 10%).

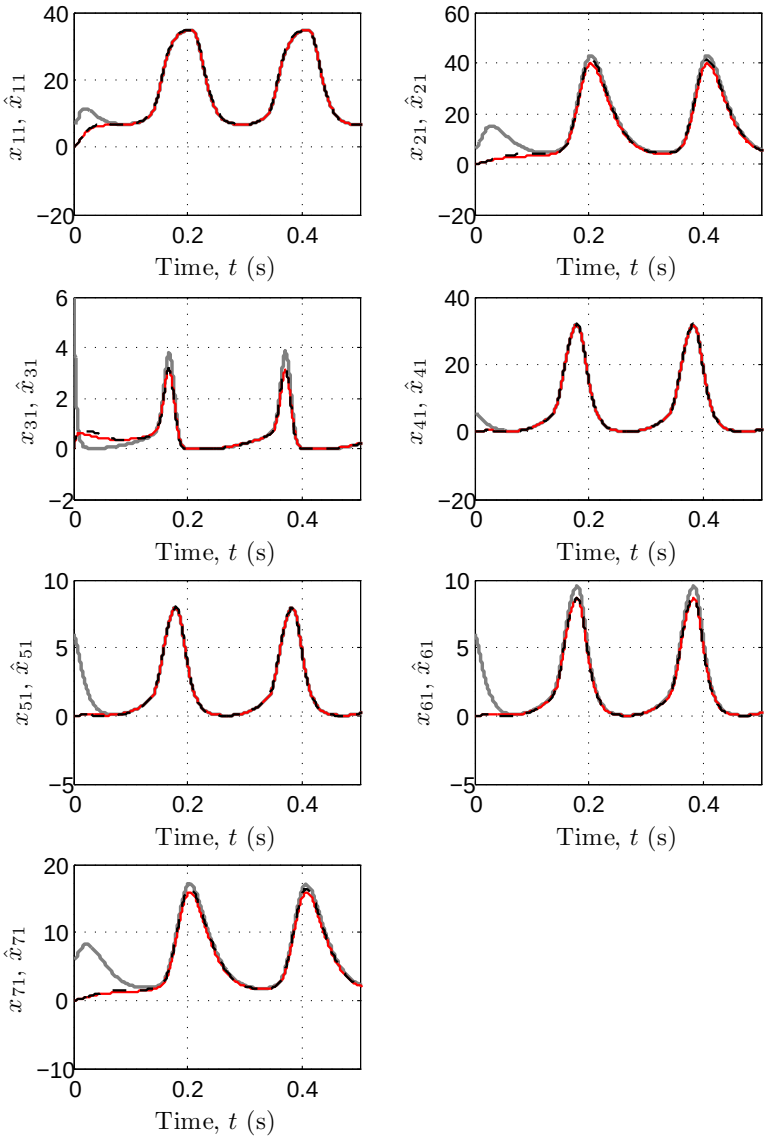


Figure 3.4: True membrane potential contribution x_{i1} (grey solid line) and estimated membrane potential contribution $\hat{x}_{i1}$ (red solid line: $l = k = 0$, black dashed line: $k = 0.1$, $l = -0.5$) under the practical scenario (Assumptions 5 to 8).

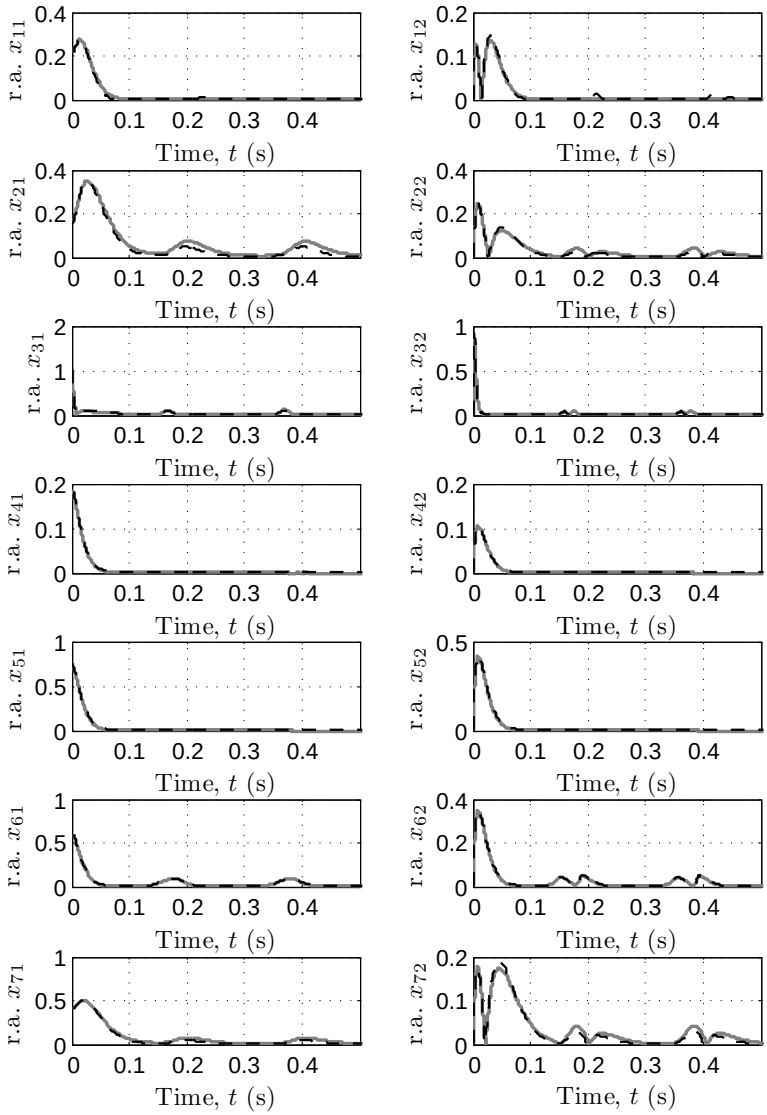


Figure 3.5: Relative state estimation error $\frac{|x_i - \hat{x}_i|}{\max(x_i) - \min(x_i)}$ for $i \in \{1, \ldots, 7\}$ under the practical scenario (Assumptions 5 to 8) for the estimator (grey solid line) and the observer with $k = 0.1$, $l = -0.5$ (black dashed line).

3.3.3 Simulations under Assumptions 1, 3, 4 and 6: uncertain-input observer.

The input to the brain is usually hard to quantify under realistic conditions. To gauge the performance of the observers under this condition, we perform another set of simulations where the input is set to $u = 0$ for the observers with no uncertainty in parameters and measurement noise. This case is formally stated in Remark 3.

For both observers, Figures 3.7 to 3.6 illustrate that the estimation error converges to a neighbourhood of the origin as expected. In Figure 3.7, the estimation error for subsystems (x_{21}, x_{22}), (x_{31}, x_{32}), (x_{41}, x_{42}), (x_{51}, x_{52}), (x_{61}, x_{62}) and (x_{71}, x_{72}) converges to a reasonably small neighbourhood of the origin. However, the (x_{11}, x_{12}) subsystem exhibits a much larger steady state estimation error than the other subsystems due to the input u directly affecting it. Nevertheless, our simulations in Figure 3.8 show that the states can be reasonably reconstructed when the input is unknown.

In Figure 3.6, the estimator outperforms the observer with $k = 0.1$ and $l = -0.5$, in the sense that the estimator converges to a smaller neighbourhood around the origin $\tilde{x} = 0$ (with $|\tilde{x}| \leq 35$ for the estimator and $|\tilde{x}| \leq 85$ for the observer with $l = -0.5$ and $k = 0.1$). A decision regarding the tradeoff between the convergence speed of the state estimation error and the accuracy of the estimates needs to be made when employing these observers. In this simulation scenario, the error converges to a neighbourhood in a reasonable timeframe ($t = 0.2$s) for both observers. Hence, the estimator is the observer of choice due to the steady-state accuracy of its estimates.

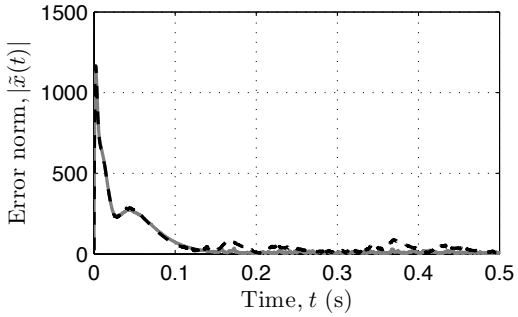


Figure 3.6: Norm of state estimation error $|\tilde{x}|$ for the estimator (grey solid line) and the observer with $k = 0.1$, $l = -0.5$ (black dashed line) under Assumptions 1, 3, 4 and 6.

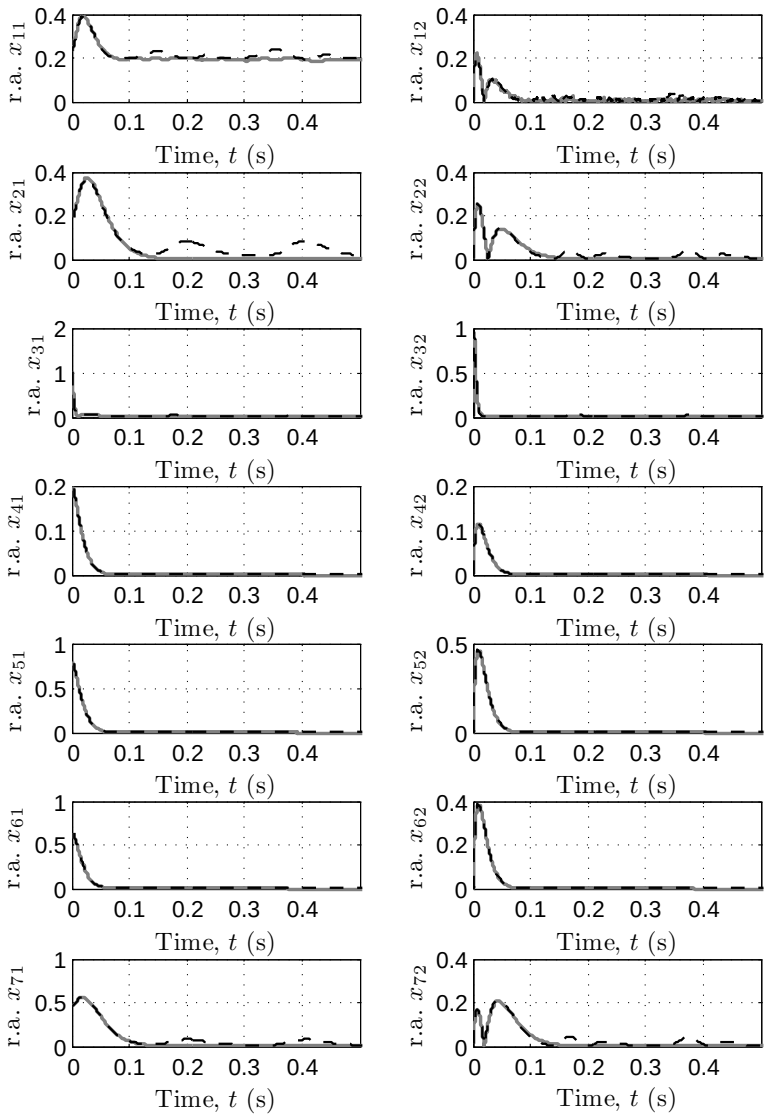


Figure 3.7: Absolute state estimation error $\frac{|x_i - \hat{x}_i|}{\max(x_i) - \min(x_i)}$ for $i \in \{1, \ldots, 7\}$ for the estimator (grey solid line) and the observer with $k = 0.1$, $l = -0.5$ (black dashed line) under Assumptions 1, 3, 4 and 6.

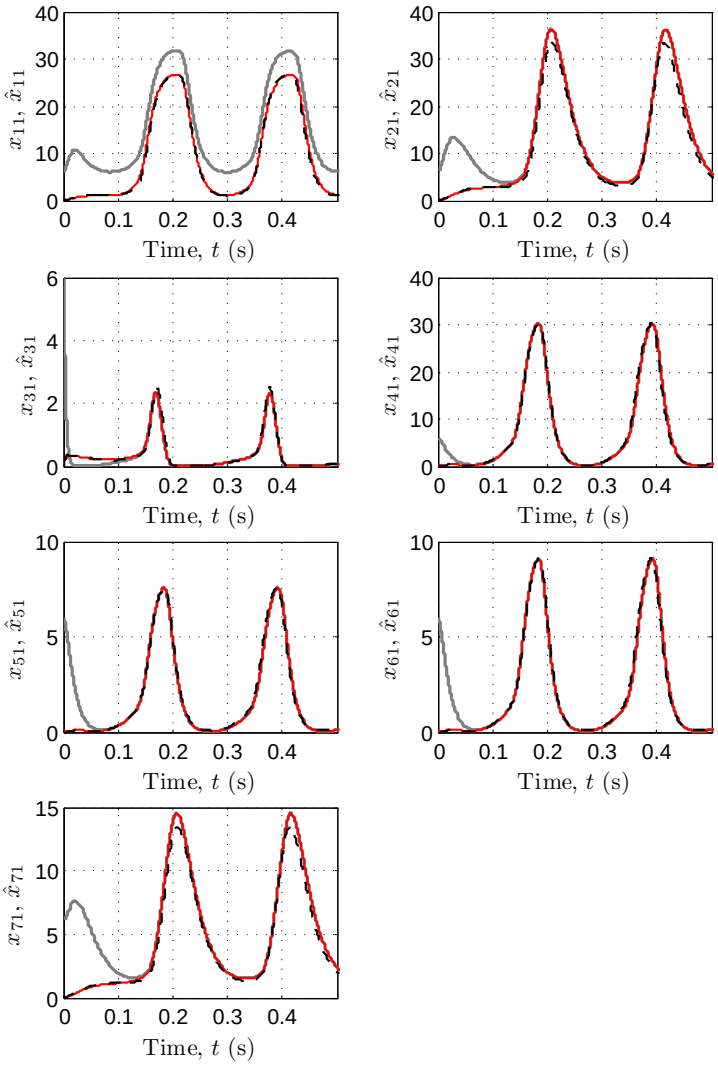


Figure 3.8: True membrane potential contribution x_{i1} (grey solid line) and estimated membrane potential contribution $\hat{x}_{i1}$ (red solid line: $k = l = 0$, black dashed line: $k = 0.1$, $l = -0.5$) under Assumptions 1, 3, 4 and 6.

3.4 Summary

A nonlinear observer is proposed for the class of neural mass models described in Chapter 2. The observer is a global exponential observer (see Definition 1 in Part I), i.e. the state estimates converge to the true states for all initial conditions in exponential time. We showed a posteriori that the proposed observer is robust towards disturbances and uncertainties. Simulation results for the model by Wendling et. al. [155] confirm our main results.

The reader is reminded that the nonlinear observer designed in this chapter is specific to the class of neural mass models, due to the cascade property of the observation error systems that enables us to show the global exponentially convergent property of the nonlinear observer. In the subsequent chapter, the circle criterion observers presented is not model specific, but applicable to systems that can be written in the desired form (2.18), with a nonlinearity γ that satisfies certain assumptions.

Chapter 4

A robust circle criterion observer

In this chapter, we propose to apply circle criterion observers which have been originally developed by Arcak and Kokotović in [16] and extended in [17; 45; 158]. Unfortunately, none of these designs apply to the neural mass models studied because the required linear matrix inequality (LMI) condition is not feasible. Hence, we combine the results for monotonic and globally Lipschitz nonlinearities [17; 158], as well as introduce a multiplier [45] to obtain a less restrictive LMI that is feasible for the neural mass models considered. This first result is presented in Section 4.1.

We also address two main issues faced in neuroscientific studies. Firstly, the input is not always measurable. Secondly, the measurements obtained are corrupted by noise. Hence, we improve the observer design in [35] by taking into account these two implementation issues. The resulting design allows observer gain matrices L and K to be obtained under the circle criterion, while taking the attenuation of input uncertainty and measurement noise into account. Our design differs from [158] in that we consider input uncertainty and we also introduce a multiplier M in the LMI, so that the resulting observer is applicable to the class of neural mass models we consider. We present this extension in Section 4.2

4.1 A circle criterion observer

This section presents the first extension of the circle criterion observers in [16; 17; 45; 158] such that they are applicable to the neural mass models considered in Chapter 2. *A posteriori* analysis of the robustness of the extended circle criterion observer with respect to uncertainty and disturbances is carried out in Section 4.1.1. We then show how the observer can be applied to neural mass models in Section 4.1.2 and provide

simulations in Section 4.1.3 to illustrate our results.

Recall that we consider the class of nonlinear systems stated in (2.8) and reproduced here

$$\begin{aligned} \dot{x} &= Ax + G(p^\star)\gamma(Hx) + \sigma(u, y, p^\star) \\ y &= Cx, \end{aligned} \tag{4.1}$$

where the state vector is $x \in \mathbb{R}^{n_x}$, the input is $u \in \mathbb{R}^{n_u}$, the measurement is $y \in \mathbb{R}^{n_y}$, $\gamma = (\gamma_1, \ldots, \gamma_m) : \mathbb{R}^{n_m} \to \mathbb{R}^{n_m}$ and $\sigma = (\sigma_1, \ldots, \sigma_n) : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_p} \to \mathbb{R}^{n_x}$. Suppose γ is both globally Lipschitz and monotonically increasing as follows

Assumption 9. *For any $i \in \{1, \ldots, n_m\}$, there exists constants $0 \leq a_i \leq b_i < \infty$, so that the following holds:*

$$a_i \leq \tfrac{\gamma_i(v_i) - \gamma_i(w_i)}{v_i - w_i} \leq b_i, \ \forall v_i, w_i \in \mathbb{R} \text{ with } v_i \neq w_i.$$

Assumption 9 is an extension of the slope restriction condition from [17, Equation 1] to vector nonlinearity $\gamma = (\gamma_1, \ldots, \gamma_m)$. Constant b_i is the Lipschitz constant of γ_i.

We consider the following type of observer originally proposed in [16]

$$\dot{\hat{x}} = A\hat{x} + G(p^\star)\gamma\big(H\hat{x} + K(C\hat{x} - y)\big) + L(C\hat{x} - y) + \sigma(u, y, p^\star), \tag{4.2}$$

where $\hat{x}$ is the state estimate and K, L are observer matrices to be designed.

Denoting the observation error as $\tilde{x} := x - \hat{x}$ and $\eta := v - w$ where $v := Hx$ and $w := H\hat{x} + K(C\hat{x} - y)$, the observation error system from (4.1) and (4.2) is

$$\dot{\tilde{x}} = (A + LC)\tilde{x} + G(p^\star)\big(\gamma(v) - \gamma(w)\big).$$

Note that from Assumption 9, we know that for any $i \in \{1, \ldots, m\}$, there exists a time-varying gain $\delta_i(t)$ taking values in the interval $[0, b_i]$ so that

$$\gamma_i(v_i(t)) - \gamma_i(w_i(t)) = \delta_i(t)(v_i(t) - w_i(t)), \ \forall v_i, w_i \in \mathbb{R}. \tag{4.3}$$

By (4.3), we obtain the observation error system as

$$\dot{\tilde{x}} = (A + LC)\tilde{x} + G(p^\star)\delta(t)\eta, \tag{4.4}$$

where $\delta(t) = \operatorname{diag}(\delta_1(t), \ldots, \delta_m(t))$.

We show in Theorem 3 that the origin of the observation error system (4.4) is

globally exponentially stable (GES) under certain conditions.

Theorem 3. *Suppose $x(t)$ exists for all $t \geq 0$. Under Assumptions 1-3 and 9, if there exist a matrix $P = P^T > 0$, a diagonal matrix $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_m)$ with strictly positive components and a constant $\nu > 0$ such that*

$$\begin{bmatrix} (A+LC)^T P + P(A+LC) + \nu I & PG(p^\star) + (H+KC)^T \Lambda \\ G(p^\star)^T P + \Lambda(H+KC) & -2\Lambda diag\left(\frac{1}{b_1}, \ldots, \frac{1}{b_m}\right) \end{bmatrix} \leq 0, \tag{4.5}$$

then there exists k, $\beta > 0$ such that the following holds:

$$|\tilde{x}(t)| \leq k \exp(-\beta t)|\tilde{x}(0)|, \qquad \forall t \geq 0, \forall \tilde{x}(0) \in \mathbb{R}^{n_x}. \tag{4.6}$$

□

Theorem 3 states that an observer of the form (4.2) can be designed for system (4.1) if the observer matrices K and L can be found such that LMI (4.5) is satisfied. As explained in [16] and later in Section 4.1.2, inequality (4.5) can be treated as an LMI in P, PL, Λ, ΛK and ν. Hence, widely available software packages such as the LMI Lab in MATLAB can be used to solve (4.5).

Existing circle criterion observer results [17; 45; 158] yield LMIs that are infeasible for the neural models we consider in Chapter 2. Therefore, we tailored the results of [45] to the case where vector nonlinearities γ are not only monotonically increasing but also globally Lipschitz. In that way, the LMI condition (4.5) differs from the LMI in (17) of [45], where element $(2,2)$ in (4.5) is non-zero which makes it less restrictive. The LMI condition (4.5) also differs from (13) of [158], where the presence of multiplier Λ in elements $(1,2)$, $(2,1)$ and $(2,2)$ of (4.5), gives us more flexibility. Hence, we are able to design circle criterion observers for the neural models of interest.

4.1.1 Robustness analysis

The robustness of the observer towards uncertainties and disturbances is important for its practicality in a realistic setting. Therefore, we introduce ϵ_y for measurement noise (Assumption 5), ϵ_u for input uncertainty (Assumption 6) and ϵ_{sys} to characterise modelling uncertainty (Assumption 8). Instead of (4.1) and (4.2), our model and

observer are now restated as follows

$$\begin{aligned} \dot{x} &= Ax + G(p^\star)\gamma(Hx) + \sigma(u, y, p^\star) + \epsilon_{sys} \\ y &= Cx, \qquad (4.7) \\ \dot{\hat{x}} &= A\hat{x} + G(p^\star)\gamma\Big(H\hat{x} + K\big(C\hat{x} - (y + \epsilon_y)\big)\Big) \\ &\quad + L\big(C\hat{x} - (y + \epsilon_y)\big) + \sigma(u + \epsilon_u, y + \epsilon_y, p^\star). \qquad (4.8) \end{aligned}$$

We make the following additional assumption:

Assumption 10. *Each component* $\sigma_i : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_p} \to \mathbb{R}$ *for* $i \in \{1, \ldots, n_x\}$ *is globally Lipschitz with constant* $\tilde{b}_i$.

Note that due to Assumption 9-10 and assuming that $\epsilon_{sys} \in \mathcal{L}_\infty$, the solution $x(t)$ exists for all time $t \geq 0$ [85, Theorem 3.2]. The following theorem shows that the observation error $\tilde{x}$ from (4.7) and (4.8) is input-to-state stable (ISS) [131] with respect to disturbances ϵ_y, ϵ_u and uncertainty ϵ_{sys}.

Theorem 4. *Consider the perturbed model* (4.7) *and observer* (4.8). *If Assumptions 9-10 hold and LMI* (4.5) *is feasible, then there exist* $\tilde{k}$, $\tilde{\beta}$, γ_u, γ_y, $\gamma_{\epsilon_{sys}} > 0$ *such that* (4.9) *holds for all* $t \geq 0$, $\tilde{x}(0) \in \mathbb{R}^{n_x}$ *and for all* ϵ_u, ϵ_y, $\epsilon_{sys} \in \mathcal{L}_\infty$,

$$|\tilde{x}(t)| \leq \tilde{k}\exp(-\tilde{\beta}t)|\tilde{x}(0)| + \gamma_u\|\epsilon_u\|_{[0,t]} + \gamma_y\|\epsilon_y\|_{[0,t]} + \gamma_{\epsilon_{sys}}\|\epsilon_{sys}\|_{[0,t]}. \qquad (4.9)$$

□

The ISS property guarantees that the observation error converges to a neighbourhood of the origin whose size depends on the norms of input uncertainty ϵ_u, modelling uncertainty ϵ_{sys} and disturbance in the measurement ϵ_y. Note that we recover the global exponential property of the observation error system in Theorem 3 when all the uncertainties and disturbances are set to 0.

4.1.2 Application to the model by Jansen and Rit

In this section, we show how Theorem 4 can be applied to the class of neural mass models introduced in Chapter 2. We consider the model by Jansen and Rit described in Chapter 2.3.2 and written in the form of (4.1) in Chapter 2.4.3.

Under practical conditions, the parameters θ_A, θ_B, C_1, C_2, C_3 and C_4 are not known and are time-varying (see Assumption 5 in Chapter 3.2). To characterise the

uncertainty in the parameters, we introduce ϵ_p so that the true parameter of the system (4.1) is $p^\star + \epsilon_p$, where ϵ_p is bounded and unknown. Therefore, the model (4.1) becomes:

$$\begin{aligned} \dot{x} &= Ax + G(p^\star + \epsilon_p)\gamma(Hx) + \tilde{\sigma}(u, y, p^\star + \epsilon_p), \\ y &= Cx. \end{aligned} \tag{4.10}$$

Moreover, the EEG measurement y is not noise-free and the input u from afferent populations is not quantifiable in practice. As performed in Section 4.1.1, we introduce ϵ_y to characterise measurement noise such that the EEG measurement available to the observer is now $y + \epsilon_y$ and the input u is allowed to be uncertain by introducing ϵ_u such that the input available to the observer is $u + \epsilon_u$.

We consider the observer in the form of (4.8) as follows:

$$\begin{aligned} \dot{\hat{x}} &= A\hat{x} + G(p^\star)\gamma\Big(H\hat{x} + K(p^\star)\big(C\hat{x} - (y + \epsilon_y)\big)\Big) \\ &\quad + \sigma(u + \epsilon_u, y + \epsilon_y, p^\star) + L(p^\star)(C\hat{x} - (y + \epsilon_y)), \end{aligned} \tag{4.11}$$

where $K(p^\star)$ and $L(p^\star)$ are obtained by solving LMI (4.5).

Note that as the LMI (4.5) is dependent on $G(p^\star)$, verifying that (4.5) is satisfied for $p^\star \in \Theta := [\theta_{A_{min}}, \theta_{A_{max}}] \times [\theta_{B_{min}}, \theta_{B_{max}}] \times [C_{1_{min}}, C_{1_{max}}] \times [C_{2_{min}}, C_{2_{max}}] \times [C_{3_{min}}, C_{3_{max}}] \times [C_{4_{min}}, C_{4_{max}}]$ will require infinite number of LMIs to be checked. To avoid this, we proceed as follows. We note that $G(p^\star)$ is:

$$G(p^\star) = (0, \theta_1 a, 0, \theta_2 b, 0, 0, 0, 0), \tag{4.12}$$

where $\tilde{\theta} := (\theta_1, \theta_2) = (\theta_A C_2, \theta_B C_4)$. Henceforth, we denote $G(p^\star)$ as $G(\tilde{\theta})$ with slight abuse of notation. The tuple of parameters (θ_1, θ_2) takes values in the convex set $\tilde{\Theta} := [\underline{\theta}_1, \bar{\theta}_1] \times [\underline{\theta}_2, \bar{\theta}_2]$, where $\bar{\theta}_1 = \theta_{A_{max}} C_{2_{max}}$, $\underline{\theta}_1 = \theta_{A_{min}} C_{2_{min}}$, $\bar{\theta}_2 = \theta_{B_{max}} C_{4_{max}}$ and $\underline{\theta}_2 = \theta_{B_{min}} C_{4_{min}}$. The vertices of $\tilde{\Theta}$ are then:

$$\begin{cases} \theta_1^\star = (\underline{\theta}_1, \underline{\theta}_2) \\ \theta_2^\star = (\underline{\theta}_1, \bar{\theta}_2) \\ \theta_3^\star = (\bar{\theta}_1, \bar{\theta}_2) \\ \theta_4^\star = (\bar{\theta}_1, \underline{\theta}_2). \end{cases} \tag{4.13}$$

We now show that if LMI (4.5) is solvable for each $\theta_i^\star \in \tilde{\Theta}$, $i \in \{1, \ldots, 4\}$ with the same positive definite matrix $P = P^T$, then it is solvable for any $p^\star \in \Theta$. In that way, only four LMIs need to be verified to ensure the feasibility of (4.5). Due to the convexity of

$\tilde{\Theta}$, any point $\tilde{\theta} = (\theta_1, \theta_2) \in \tilde{\Theta}$ can be written in the form of:

$$\tilde{\theta} = \sum_{i=1}^{4} \lambda_i \theta_i^\star \tag{4.14}$$

where $\lambda_i \geq 0$, $i \in \{1, \ldots, 4\}$ and $\sum_{i=1}^{4} \lambda_i = 1$. From (4.12), $G(\tilde{\theta})$ is linear in $\tilde{\theta}$, hence:

$$G(\tilde{\theta}) = \sum_{i=1}^{4} \lambda_i G(\theta_i^\star). \tag{4.15}$$

Now, we convert the matrix inequality of Theorem 3 into the following LMI as explained in [16]:

$$\begin{aligned} &\mathcal{M}_{\tilde{\theta}}(P, R, S, \Lambda, \nu) = \\ &\begin{bmatrix} A^{\mathrm{T}}P + PA + RC + C^{\mathrm{T}}R^{\mathrm{T}} + \nu\mathbb{I} & PG(p^\star) + H^{\mathrm{T}}\Lambda + C^{\mathrm{T}}S^{\mathrm{T}} \\ G(p^\star)^{\mathrm{T}}P + \Lambda H + SC & -2\Lambda \mathrm{diag}(b_1^{-1}, \ldots, b_m^{-1}) \end{bmatrix} \leq 0 \end{aligned} \tag{4.16}$$

where $R = PL$ and $S = \Lambda K$. Suppose that there exist positive definite matrix $P = P^T$, matrices R_i and S_i, positive definite diagonal matrix Λ_i and strictly positive constant ν_i for $i \in \{1, \ldots, 4\}$ such that the following holds:

$$\mathcal{M}_{\theta_i^\star}(P, R_i, S_i, \Lambda_i, \nu_i) \leq 0, \tag{4.17}$$

for each vertex $\theta_i^\star$ defined in (4.13). Consider $\tilde{\theta} \in \tilde{\Theta}$, we know that there exist λ_i such that (4.14) holds for all $\lambda_i \geq 0$ and $\sum_{i=1}^{4} \lambda_i = 1$. Due to the linearity of $\mathcal{M}_\theta(P, R, S, \Lambda, \nu)$ in its arguments and (4.15), the following is satisfied:

$$\mathcal{M}_{\tilde{\theta}}(P, R, S, \Lambda, \nu) = \sum_{i=1}^{4} \lambda_i \mathcal{M}_{\theta_i^\star}(P, R_i, S_i, \Lambda_i, \nu_i), \tag{4.18}$$

where $R = \sum_{i=1}^{4} \lambda_i R_i$, $S = \sum_{i=1}^{4} \lambda_i S_i$, $\Lambda = \sum_{i=1}^{4} \lambda_i \Lambda_i$ is a positive definite diagonal matrix and $\nu = \sum_{i=1}^{4} \lambda_i \nu_i > 0$. Consequently, the feasibility of all the LMIs (4.17) for all $i \in \{1, \ldots, 4\}$ with the same P implies that the LMI (4.5) (equivalently (4.16)) is satisfied for all $\theta \in \Theta$. Note that the same P that satisfies LMI (4.16) also has to work for all LMIs (4.17) for $i \in \{1, \ldots, 4\}$, i.e. we have a simultaneous Lyapunov function for all vertices [85, Section 10.1.3]. We have computationally verified that the LMIs (4.17) are satisfied for the numerical values provided in [74, Section 2.3]. Therefore, we are able to state the following result which guarantees the existence of observer matrices

$K(p^\star)$ and $L(p^\star)$ for all $p^\star \in \Theta$ such that the state estimation error system $\tilde{x}$ is ISS with respect to the input and parameter uncertainties ϵ_u, ϵ_p as well as measurement noise ϵ_y.

Proposition 1. *Consider the perturbed model (4.10) and observer (4.11). For any $p^\star \in \Theta$ (where Θ is a compact set), there exist observer matrices $L(p^\star)$ and $K(p^\star)$ that satisfy LMI (4.5) such that the observation error system $\tilde{x}$ satisfies for all $t \geq 0$, $\tilde{x}(0) \in \mathbb{R}^8$ and for all ϵ_u, ϵ_y, $\epsilon_p \in \mathcal{L}_\infty$,*

$$|\tilde{x}(t)| \leq \bar{k}\exp(-\bar{\beta}t)|\tilde{x}(0)| + \bar{\gamma}_y\|\epsilon_y\|_{[0,t]} + \bar{\gamma}_u\|\epsilon_u\|_{[0,t]} + \bar{\gamma}_p\|\epsilon_p\|_{[0,t]}, \tag{4.19}$$

where $\bar{k}$, $\bar{\beta}$, $\bar{\gamma}_y$, $\bar{\gamma}_u$, $\bar{\gamma}_p \geq 0$. □

Similar results can be derived for the models by Stam et. al. [138] (described in Chapter 2.3.3) and Wendling et. al .[155] (described in Chapter 2.3.1) respectively. We have verified that the conditions of Theorem 4 and Proposition 1 are met.

4.1.3 Simulation results

We show the simulation results for the model (4.10) and observer (4.8). The simulations are performed under two scenarios: (1) without disturbances and uncertainties, (2) under the more practical condition with disturbances and uncertainties.

We initialise the model at $x(0) = (6, 6, 6, 6, 6, 6, 6, 6)$ and the observer at $\hat{x}(0) = (0, 0, 0, 0, 0, 0, 0, 0)$. The input to the model u as described in [74] is uniformly distributed between 120 and 320mV. The parameters were chosen to correspond to alpha-like activity $p^\star = (\theta_A, \theta_B, C_1, C_2, C_3, C_4) = (3.25, 22, 135{,}108, 33.75, 33.75)$ as identified in [74]. All other constants used in simulation are as described in [74, Section 3.1].

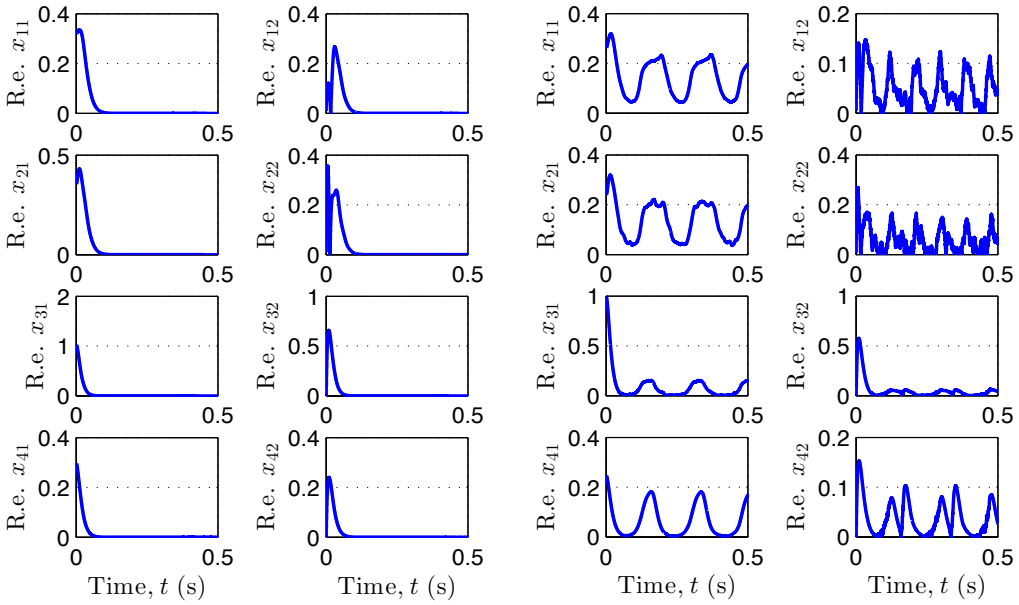
By solving the LMI (4.5), we obtain the observer matrices $L = 10^4 \times (0.0053, -2.2306, 0.0077, 5.3849, 0.0032, -0.1266, -0.0017, 0.0514)$ and $K = (-0.0586, -0.1422)$.

Scenario 1: the ideal condition without disturbance and uncertainties

In the ideal scenario, Theorem 3 states that the observer provides estimates that converge to the true states in exponential time. As illustrated in Figure 4.1(a), at $t = 0.5$s, the absolute observation error is less than 0.01%.

Scenario 2: robustness of the observer towards uncertainty in parameters, input and measurement

We now introduce parameter uncertainty that is 10% of the actual values $\epsilon_p = 0.1 \times (\theta_A, \theta_B, C_1, C_2, C_3, C_4) = (0.33, 2.2, 13.5, 10.8, 3.375, 3.375)$, input uncertainty $\epsilon_u \sim \mathcal{N}(0, 10^2)$ and measurement noise $\epsilon_y \sim \mathcal{N}(0, 0.7^2)$, as shown in our perturbed model (4.10) and observer (4.8). Our simulation results (Figure 4.1(b)) confirm the results stated in Proposition 1, i.e. the norm of the observation error is small with small $\mathcal{L}_\infty$ norm of the uncertainties. The absolute observation error relative to the amplitude of the signal as shown in Figure 4.1(b) remains less than 25% for all states.



(a) Scenario 1: the ideal condition without disturbance and uncertainties.

(b) Scenario 2: robustness of the observer towards uncertainty in parameters, input and measurement

Figure 4.1: Absolute observation error relative to the amplitude of the signal, $\frac{|\tilde{x}_i|}{|\max(x_i) - \min(x_i)|}$, for $i \in \{1, \ldots, 4\}$.

4.2 A *robust* circle criterion observer

In the previous section, we analysed the robustness of the designed observer *a posteriori* and showed that it is robust towards disturbances and uncertainty. In this section, we improve the circle criterion observer designed in Section 4.1 such that it is *a priori*

robust towards measurement noise and input uncertainty by providing the user with the ability to attenuate the effect of noise and uncertainty. To this end, we take into account measurement noise (Assumption 7) when considering the class of systems

$$\begin{aligned} \dot{x} &= Ax + G(p^\star)\gamma(Hx) + Bu \\ y &= Cx + Dw, \end{aligned} \tag{4.20}$$

where $w \in \mathbb{R}^{n_w}$ is the measurement noise. We make the same assumptions about γ as done previously in Section 4.1.

We consider the following type of observer [17]

$$\dot{\hat{x}} = A\hat{x} + G\gamma\big(H\hat{x} + K(C\hat{x} - y)\big) + L(C\hat{x} - y) + B(u + d), \tag{4.21}$$

where $\hat{x}$ is the state estimate, $d \in \mathbb{R}^{n_d}$ is the input disturbance and K, L are the observer matrices to be designed.

As done in [17], denoting the observation error as $\tilde{x} := \hat{x} - x$, $v := Hx$ and $z := H\hat{x} + K(C\hat{x} - y)$, the observation error system from (4.20) and (4.21) is $\dot{\tilde{x}} = (A + LC)\tilde{x} - LDw + Bd + G\big(\gamma(z) - \gamma(v)\big)$. By (4.3), we obtain the observation error system as

$$\dot{\tilde{x}} = (A + LC)\tilde{x} - LDw + Bd + G\delta(t)\eta, \tag{4.22}$$

where $\delta(t) = \operatorname{diag}(\delta_1(t), \ldots, \delta_m(t))$ and $\eta := z - v$.

Given the observation error system (4.22), our task is to find observer matrices K and L such that a quadratic Lyapunov function $V(\tilde{x})$ satisfies the following along the solutions of (4.22)

$$\dot{V}(\tilde{x}) \leq -|\tilde{x}|^2 + \mu_w|w|^2 + \mu_d|d|^2. \tag{4.23}$$

We can then show that the observation error $\tilde{x}$ satisfies the following property[1] for all $t \geq 0$

$$\|\tilde{x}(t)\|_2 \leq \bar{c}|\tilde{x}(0)| + \sqrt{\mu_w}\|w(t)\|_2 + \sqrt{\mu_d}\|d(t)\|_2, \tag{4.24}$$

where scalars $\bar{c}, \mu_w, \mu_d > 0$. The disturbance gains from w and d to $\tilde{x}$ are $\sqrt{\mu_w}$ and $\sqrt{\mu_d}$ respectively.

[1]We can obtain (4.24) from (4.23) by following the same procedure as in the proof of Theorem 5.2 in [86].

4.2.1 Main result

In Theorem 5, we establish that the observation error system (4.22) satisfies property (4.24) provided that a linear matrix inequality (LMI) is satisfied.

Theorem 5. *Consider system (4.20) and observer (4.21). Under Assumption 9, if there exist a real symmetric and positive definite matrix P, a diagonal and positive definite matrix $M = diag(m_1, \ldots, m_{n_m})$, and scalar constants $\mu_w, \mu_d > 0$, such that the following is satisfied*

$$\begin{bmatrix} \mathcal{A}(P, PL) & \mathcal{B}(P, M, K^T M) & -PLD & PB \\ \star & \mathcal{E}(M) & -MKD & 0 \\ \star & \star & -\mu_w \mathbb{I} & 0 \\ \star & \star & \star & -\mu_d \mathbb{I} \end{bmatrix} \leq 0, \tag{4.25}$$

where $\mathcal{A}(P, PL) = P(A+LC)+(A+LC)^T P+\mathbb{I}$, $\mathcal{B}(P, M, K^T M) = PG+(H+KC)^T M$ and $\mathcal{E}(M) = -2M\, diag\left(\frac{1}{b_1}, \ldots, \frac{1}{b_{n_m}}\right)$, then the observation error system satisfies property (4.24). □

The proof of Theorem 5 is provided in Appendix B.6. Theorem 5 shows that if a K and L can be found such that the LMI (4.25) is satisfied, then an observer (4.21) can be designed for system (4.20). Note that condition (4.25) is considered an LMI in P, PL, MK, M, μ_w and μ_d. As such, (4.25) can be solved using efficient software tools such as the LMI Lab in MATLAB.

By considering the system (4.20) under the ideal condition where there is no input uncertainty and measurement error as done in Section 4.1, we obtain the condition stated in Theorem 3. Current circle criterion results in [17], [45], [158] yield LMIs that are not feasible for the class of neural models we consider. Therefore, we adapted [45] to the case where the nonlinearity γ is globally Lipschitz and also monotonically increasing with inspiration from [17]. This result is a special case of the system considered in Theorem 5 as stated in Theorem 3 and was reported in [35]. In this section, we further improve the circle criterion observer obtained in [35] by designing observer matrices K and L under the circle criterion condition and taking input uncertainty d from (4.21) and measurement noise w from (4.20) into account. The LMI (4.25) differs from (13) obtained in [158] in the sense that we consider input uncertainty attenuation and introduced a multiplier M in components $(1, 2)$, $(2, 2)$ and $(2, 3)$ of (4.25). Without introducing the multiplier M, the results obtained in [158] do not lead to feasible LMIs. This simple extension allows circle criterion observers to be designed for the neural

models we consider, in addition to taking into account the realistic issues faced when implementing these observers in the context of estimation for neuroscientific studies.

The constants μ_w and μ_d in (4.25) may be specified by the user and should the LMI (4.25) be found to be solvable, we then have the estimation error satisfying property (4.24) with estimates of the disturbance gains $\sqrt{\mu_w}$ and $\sqrt{\mu_d}$. In some cases, we may wish to minimise these constants and various methods are available to solve this multi-objective optimisation problem (see [42]). A simple approach that we take in the next section is to minimise the cost $J_{max} = \max\{\mu_w, \mu_d\}$ subject to (4.25).

4.2.2 Application to the model by Jansen and Rit

In this section, we show how Theorem 5 can be applied to the model by Jansen and Rit described in Chapter 2.3.2 and it can be written in the form of (4.20) with $D = 1$ and all other variables as defined in Chapter 2.4.3. We introduce input disturbance $d \sim \mathcal{N}(0, 0.1^2)$ and measurement noise $w \sim \mathcal{N}(0, 0.7^2)$. The performance of **(A)** the circle criterion observer obtained under the conditions of Theorem 3 that does not consider the attenuation of input uncertainty and measurement noise is compared with **(B)** the robust circle criterion observer derived in Theorem 5. We solved LMI (4.5) to obtain observer matrices K, L for observer **(A)**. For observer **(B)**, we choose to minimise μ_w and μ_d using the cost function J_{max} subject to (4.25) to obtain K and L. The resulting computed disturbance gains are $\sqrt{\mu_w} = 706$ and $\sqrt{\mu_d} = 9.48$. In the simulation that follows, we initialise the model at $x(0) = (6, 0.5, 6, 0.5, 6, 0.5, 6, 0.5)$ and the observers at $\hat{x}(0) = 0$. Figure 4.2 shows that the robust circle criterion observer obtained in Theorem 5 (Observer **(B)**) outperforms the observer obtained in Theorem 3 (Observer **(A)**) in the presence of input uncertainty and measurement noise.

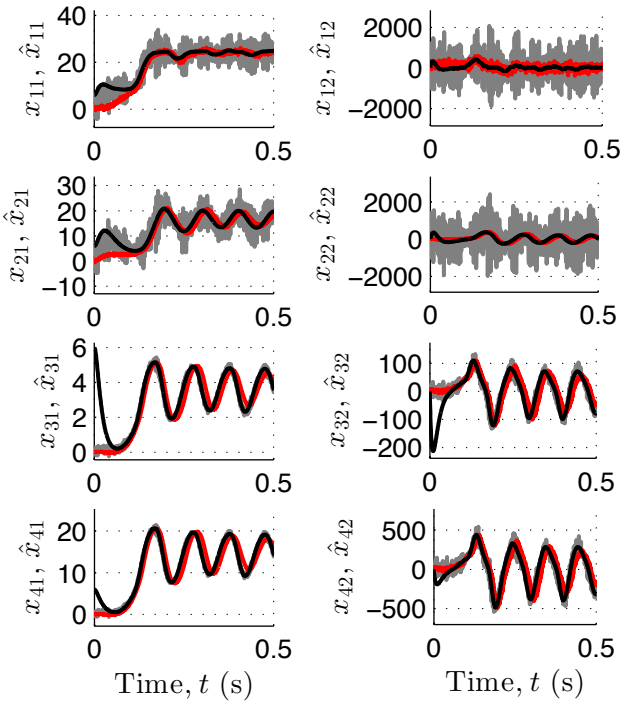


Figure 4.2: Estimated states $\hat{x}$ converge to a neighbourhood of the true states x. Legend: Observer A (grey), Observer B (red) and Model (black).

4.3 Summary

We have designed a robust circle criterion observer that attenuates input uncertainty and measurement noise. The synthesis of these observers relies on the feasibility of an LMI (see Theorems 3 and 5). To the best of our knowledge, no other results in the literature leads to feasible LMIs for the class of neural mass models we consider in Chapter 2. We extend these results to obtain feasible LMIs for our examples. As a result, we sacrifice the generality of the circle criterion observer first proposed by Arcak and Kokotović in [16], where the globally Lipschitz property of the nonlinearity γ in (4.1) is removed. The neural mass models contain a nonlinearity that is globally Lipschitz, i.e. the sigmoid function defined in (2.4) and (2.7). This provides motivation to impose a restriction on the nonlinearities (Assumption 9) as done in [17] such that the LMIs are solvable.

The circle criterion observers are applicable to any system in the class (4.1) provided the nonlinearity γ is slope-restricted. The extensions of the circle criterion observer in the literature presented in this chapter were done such that we have a *robust* nonlinear observer for the neural mass models. The robustness property of the circle criterion observers inspired the usage of these observers in a supervisory setup, developed in Chapter 6, that enables the estimation of parameters *and* states using state observers only.

Part II: Parameter and state estimation

In the previous part, we designed state observers under the assumption that the parameters p^* are known. For neuroscience applications, the assumption that the parameters are known is too restrictive and only partially useful. For the purposes of seizure monitoring and control, the parameters are unknown and need to be estimated.

Hence, in this part, we address the problem of parameter and state estimation using the setup shown in Figure 4.3 via two methods: an adaptive observer and a supervisory observer. In Chapter 5, we present an adaptive observer for a class of systems that encompasses a subset of the class of neural mass models described in Chapter 2. The parameters and states asymptotically converge to their true values for all initial conditions. In Chapter 6, we present an alternative parameter and state estimation method for *general* nonlinear systems using a multiple-model architecture usually used for control, known in the literature as supervisory control. We adapt this architecture for estimation and call it a supervisory observer. Under certain conditions, we obtain practical convergence of parameters and states.

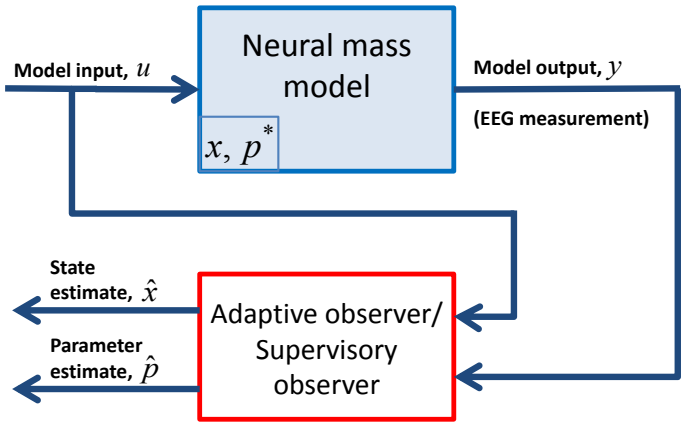


Figure 4.3: Parameter and state estimation

Chapter 5

An adaptive observer

This chapter presents an adaptive observer for a class of nonlinear systems that includes a subset of the neural mass models presented in Chapter 2. We observe that the models by Stam et al and Jansen et al described in Chapters 2.3.3 and 2.3.2 respectively admit the following general structure

$$\begin{aligned} \dot{x}_0 &= A_0 x_0 + \phi_0(y) p^\star \\ \dot{x}_1 &= A_1 x_1 + \phi_1(x_0, u) p^\star \\ y &= C_1 x_1, \end{aligned} \tag{5.1}$$

where $x_0 \in \mathbb{R}^{n_0}$, $x_1 \in \mathbb{R}^{n_1}$ are the states, $p^\star \in \mathbb{R}^{n_p}$ is a vector of constant and unknown parameters, $y \in \mathbb{R}^{n_y}$ is the measurement and $u \in \mathbb{R}^{n_u}$ is the input.

To the best of our knowledge, there is currently no adaptive observer in the literature that can be used to estimate the parameters and states of system (5.1). Indeed, most nonlinear adaptive observers apply to systems for which the nonlinearities depend on known quantities, see [161] for instance. This is not the case for system (5.1) since ϕ_1 depends on x_0 which is not measured. Few designs have been proposed for systems with state-dependent nonlinear terms, see [148] and the references therein. A notable exception is the work in [46] which focuses on a class of systems that is very similar to (5.1). While the nonlinearities in (5.1) satisfy some of the conditions stated in [46], the problem arises in the linear part which is not of the same form as in [46]. As a consequence, we need to modify the observer and the technical proof in [46]. This study extends our previous works described in Chapter 3 [33] and Chapter 4 [34; 35] on the state estimation of this class of neural mass models.

5.1 Writing the models in the form of (5.1)

In this section, we first show how a single cortical column model of neuronal populations by Jansen and Rit in [74] and described in Chapter 2.3.2 can be written in the form of (5.1). We also show that $n-$*interconnected* single cortical column models can be written in the form of (5.1). Note that the A matrix of the two models below do not satisfy the form of the A matrix in the nonlinear system considered in [46]. This reaffirms the need to design a new adaptive observer as proposed in this chapter. Our result applies to any system that can be written in the form of (5.1) and is not limited to the models described below. Nevertheless, the model by Wendling et al described in Chapter 2.3.1 cannot be written in the required form (5.1). We will discuss possible extensions of the result described in this chapter to include a more general class of nonlinear systems.

5.1.1 A single cortical column model

We are interested in estimating the membrane potential contributions of the pyramidal neurons, the excitatory and inhibitory interneurons respectively (states) and the synaptic gains of the excitatory and inhibitory interneurons (parameters). To do so, the model can be written in state space form (5.1) by taking the states to be[1] $x_0 = (x_{01}, x_{02}) \in \mathbb{R}^2$ and $x_1 = (x_{11}, \ldots, x_{14}) \in \mathbb{R}^4$, where x_{01}, x_{11}, x_{13} are membrane potential contributions of the pyramidal neurons, the excitatory and inhibitory interneurons respectively, and x_{02}, x_{12}, x_{14} are their respective derivatives. The measured output (EEG) is $y \in \mathbb{R}$, $u \in \mathbb{R}$ is the excitatory input from neighbouring columns which is assumed to be known and $p^\star = (\theta_A, \theta_B) \in \Theta \subseteq \mathbb{R}^2$ is a vector of *unknown* parameters where θ_A and θ_B represent the synaptic gains of the excitatory and inhibitory neuronal populations respectively. The matrices in (5.1) are defined as

$$C_1 = \begin{pmatrix} 1 & 0 & -1 & 0 \end{pmatrix}, \; A_0 = A_a, \; A_1 = \operatorname{diag}(A_a, A_b),$$

$$\text{where } A_a = \begin{pmatrix} 0 & 1 \\ -a^2 & -2a \end{pmatrix}, \quad A_b = \begin{pmatrix} 0 & 1 \\ -b^2 & -2b \end{pmatrix}.$$

[1] According to the notation of [74], $x_0 = (y_0, y_3)$ and $x_1 = (y_1, y_4, y_2, y_5)$.

The parameters $a, b \in \mathbb{R}_{>0}$ are assumed to be known. It has to be noted that the matrices A_0 and A_1 are Hurwitz. The nonlinear terms in (5.1) are given by

$$\begin{aligned} \phi_0(y) &= \begin{pmatrix} 0 & 0 \\ aS(y) & 0 \end{pmatrix}, \\ \phi_1(x_0, u) &= \begin{pmatrix} 0 & 0 \\ ac_2 S(c_1 x_{01}) + au & 0 \\ 0 & 0 \\ 0 & bc_4 S(c_3 x_{01}) \end{pmatrix}. \end{aligned}$$

The parameters $c_1, c_2, c_3, c_4 \in \mathbb{R}_{\geq 0}$ are assumed to be known parameters and S denotes the sigmoid function, for $v \in \mathbb{R}$: $S(v) = \frac{2e_0}{1+e^{r(v_0-v)}}$ with known constants $e_0, v_0, r \in \mathbb{R}_{>0}$. For a detailed description about the model and its parameters, see [74] or Appendix B.12.

5.1.2 Interconnected cortical column models

We now consider a model of n interconnected cortical columns presented in [154]. This model is obtained by interconnecting n of the single cortical column models described in Section 5.1.1, in the manner shown in Figure 5.1. Physiologically, the interactions between cortical columns are the firing rates of each column $S(y_i)$. Hence, a physiological interpretation of Figure 5.1 would include the sigmoid block S as part of the single column model.

The study of the interconnected models has the potential of gaining a better understanding of the interactions between different regions, whereby each model represents a part of the cortex [154]. The strength of their interaction is captured by the linear gains K_{ji}, for $i, j \in \{1, \ldots, n\}$ and $i \neq j$. The propagation dynamics of the interaction between these regions is described by:

$$h_d(t) = \theta_A a_d t \exp(-a_d t), \qquad \forall t \geq 0, \tag{5.2}$$

where θ_A, a_d are positive constants defined in Table 1 of [154].

We assume that the synaptic gains of the populations θ_{Ai} and θ_{Bi} are unknown and the coupling gains K_{ji} for $i, j \in \{1, \ldots, n\}$ and $i \neq j$ are known. The interconnected models can then be written in the form of (5.1) by taking the states in (5.1) to be[1] $x_0 = (x_{01}^1, x_{02}^1, x_{03}^1, x_{04}^1, \ldots, x_{01}^n, x_{02}^n, x_{03}^n, x_{04}^n)$ and $x_1 = (x_{11}^1, x_{12}^1, x_{13}^1, x_{14}^1, \ldots, x_{11}^n, x_{12}^n, x_{13}^n, x_{14}^n)$,

[1] According to the notation of [154] as $x_0 = (y_0^1, y_3^1, y_6^1, y_7^1, \ldots, y_0^n, y_3^n, y_6^n, y_7^n)$ and $x_1 = (y_1^1, y_4^1, y_2^1, y_5^1, \ldots, y_1^n, y_4^n, y_2^n, y_5^n)$.

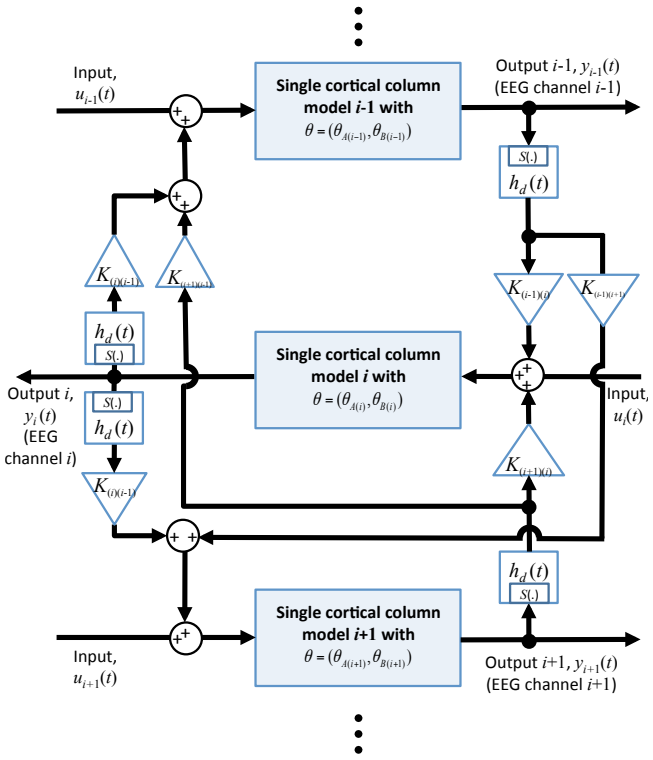


Figure 5.1: $n-$interconnected models, where $i \in \{2, \ldots, n-1\}$.

where integer $n > 1$. The input is $u = (u_1, \ldots, u_n)$ and measurement $y = (y_1, \ldots, y_n)$. The matrices are:

$$\begin{aligned} A_0 &= \text{diag}(A_{01}, \ldots, A_{0n}), \\ A_1 &= \text{diag}(A_{11}, \ldots, A_{1n}), \\ C_1 &= \begin{pmatrix} C_{11} & \ldots & C_{1n} \end{pmatrix}, \end{aligned} \tag{5.3}$$

where $C_{1i} = \begin{pmatrix} 1 & 0 & -1 & 0 \end{pmatrix}$, $A_{0i} = \text{diag}(A_a, A_d)$ and $A_{1i} = \text{diag}(A_a, A_b)$ for $i \in \{1, \ldots, n\}$, A_a and A_b are as defined in (5.2) and $A_d = \begin{pmatrix} 0 & 1 \\ -a_d^2 & -2a_d \end{pmatrix}$. The nonlinear terms in (5.1) are:

$$\begin{aligned} \phi_0(y) &= \begin{pmatrix} \phi_{01}, & \ldots, & \phi_{0n} \end{pmatrix}, \\ \phi_1(x_0, u) &= \begin{pmatrix} \phi_{11}, & \ldots, \phi_{1n} \end{pmatrix}, \end{aligned}$$

where for $i \in \{1, \ldots, n\}$,

$$\phi_{0i} = \begin{pmatrix} 0 & 0 \\ aS(y_i) & 0 \\ 0 & 0 \\ a_d S(y_i) & 0 \end{pmatrix}, \quad \phi_{1i} = \begin{pmatrix} 0 & 0 \\ au_i + ac_2 S(c_1 x_{01}^i) & \\ +\sum_{j\in\{1,\ldots,n\}, j\neq i} K_{ji} x_{03}^j & 0 \\ 0 & 0 \\ 0 & bc_4 S(c_3 x_{01}^i) \end{pmatrix}.$$

5.2 Problem formulation

For ease of notation, we write (5.1) in the following form:

$$\begin{aligned} \dot{x} &= Ax + \phi(y, u, x)p^\star \\ y &= Cx, \end{aligned} \tag{5.4}$$

where $x = (x_0, x_1)$, $A = \mathrm{diag}(A_0, A_1)$, $C = (0, C_1)$ and $\phi = (\phi_0, \phi_1)$. The nonlinear terms $\phi_0 : \mathbb{R} \to \mathbb{R}^{n_0} \times \mathbb{R}^p$ and $\phi_1 : \mathbb{R}^{n_0} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_1} \times \mathbb{R}^p$ are globally Lipschitz and bounded.

Our objective is to synthesise an adaptive observer to simultaneously estimate the state x and parameter $p^\star$ of (5.1) from the measured output y. We make the following crucial assumptions in the design of the adaptive observer:

Assumption 11. *The vector of unknown parameters θ is constant and is known to reside in a compact set Θ.*

This assumption is justified under the circumstances where the parameters are slowly-varying for each type of cortical activity [74]. When a change in activity occurs, an abrupt variation of $p^\star$ is presumed which violates Assumption 11 for only a short time as illustrated in simulations in Chapter 5.4. For the models described in Chapter 5.1, the plant's parameters $p^\star$ were identified to lie in a compact set in Sections 3.1 and 3.2 of [74] respectively.

Assumption 12. *The input u is known.*

This is a limiting assumption in practice as the input is unknown/unmeasured. A common assumption is to model the input as random white noise, see [74; 154].

Assumption 13. *The measured output y is noise-free.*

This assumption is more justified when electroencephalogram (EEG) recorded with intracranial electrodes as opposed to EEG obtained from electrodes placed on the scalp.

We investigate the robustness of our algorithm to measurement noise in simulations in Chapter 5.4.

5.3 Main result: an adaptive observer

The adaptive observer synthesised in this chapter draws upon the design by Zhang in [160] for linear systems and Farza et. al. in [46] for a class of systems that differs from the class we consider (5.1) in its linear part. Hence, we modified the observer in [46] and derive a design as presented in the following:

$$\begin{aligned}
\dot{\hat{x}} &= A\hat{x} + \phi(y,u,\hat{x})\hat{p} + \Gamma(y-\hat{y}) \\
\hat{y} &= C\hat{x} \\
\dot{\hat{p}} &= \bar{\Gamma}(y-\hat{y}) \\
\dot{\Upsilon} &= A\Upsilon + \Delta\phi(y,u,\hat{x}) \qquad \text{with } \Upsilon(0) = 0 \\
\dot{P} &= dP - dP\Upsilon^{\mathrm{T}}C^{\mathrm{T}}C\Upsilon P \qquad \text{with } P(0) = P(0)^T > 0,
\end{aligned} \tag{5.5}$$

where $\Gamma = \Delta^{-1}\Upsilon\bar{\Gamma}$, $\bar{\Gamma} = P\Upsilon^{\mathrm{T}}C^{\mathrm{T}}$ and $\Delta = \mathrm{diag}(\mathbb{I}_{n_0}, \frac{1}{d}\mathbb{I}_{n_1})$ with $d > 0$ a design parameter. The vector $\hat{x}$ denotes the estimate of x and $\hat{p}$ the estimate of $p^\star$. The variable $\Upsilon \in \mathbb{R}^{(n_0+n_1)\times n_p}$ is initialized to $\Upsilon(0) = 0$ in order to simplify the technical statements. Our result also applies when it is not the case. An essential assumption for our design to work is the condition stated below.

Assumption 14. *For any signals $u, y, \hat{x}$ that belong to $\mathcal{L}_\infty$, there exist $a_1, a_2 \in \mathbb{R}_{>0}$, $T \in \mathbb{R}_{>0}$ such that the solution to:*

$$\dot{\Upsilon} \;=\; A\Upsilon + \Delta\phi(y,u,\hat{x}) \quad \textit{with} \quad \Upsilon(0) = 0, \tag{5.6}$$

satisfies for all[1] *$t \geq 0$:*

$$a_1\mathbb{I}_{n_p} \;\leq\; \int_t^{t+T} \Upsilon^{\mathrm{T}}(\tau)C^{\mathrm{T}}C\Upsilon(\tau)\,d\tau \;\leq\; a_2\mathbb{I}_{n_p}. \tag{5.7}$$

□

The condition (5.7) is known as the persistency of excitation of the signal $C\Upsilon(t)$ and is a well-known condition in identification and adaptive control literature (see [71]).

[1] A unique solution exists for all time, for all the ordinary differential equations in (5.4) and (5.5). By Theorem 3.3 of [86], $\dot{P} = dP - dP\Upsilon^T C^T C\Upsilon P$ in (5.5) has a unique solution because its right hand side is locally Lipschitz and piecewise continuous in t and all its solution are in a compact set (by Lemma 1 of [162], P is bounded). The other ODEs in (5.4) and (5.5) have unique solutions because the nonlinearity ϕ is globally Lipschitz.

This is a sufficient condition for the identification of the parameter $p^\star$ using the adaptive observer proposed here and it is similar to the condition used in (A4) of [46] and (A3) in [24]. Inequality (5.7) is hard to verify analytically in general. Nevertheless, we have observed in simulations that this condition is satisfied for the model we consider under the simulation conditions stated in Chapter 5.4.

The condition (5.7) is needed to ensure that the solution to $\dot{P} = dP - dP\Upsilon^T C^T C \Upsilon P$ in (5.5) is symmetric and positive definite for all positive time and with lower and upper bound which are independent of d, see Lemma 1 in [162]. This is a crucial property because the matrix P will be used to construct a Lyapunov function in the proof.

We are now ready to state the main result which ensures the asymptotic convergence of the estimated variables $(\hat{x}, \hat{p})$ to $(x, p^\star)$. Its proof is given in the Appendix.

Theorem 6. *Consider the system (5.4) and the observer (5.5) and suppose Assumptions 11-14 are satisfied. Then, there exists $d^\star \geq 1$ such that for all $d \geq d^\star$, the estimates $(\hat{x}, \hat{p})$ asymptotically converge towards $(x, p^\star)$ i.e. for any $d \geq d^\star$, there exist $\beta_d \in \mathcal{KL}$ such that for any input u and any initial conditions $P(0) = P(0)^{\mathrm{T}} > 0$, $\tilde{x}(0)$ and $\tilde{p}(0)$, the following holds :*

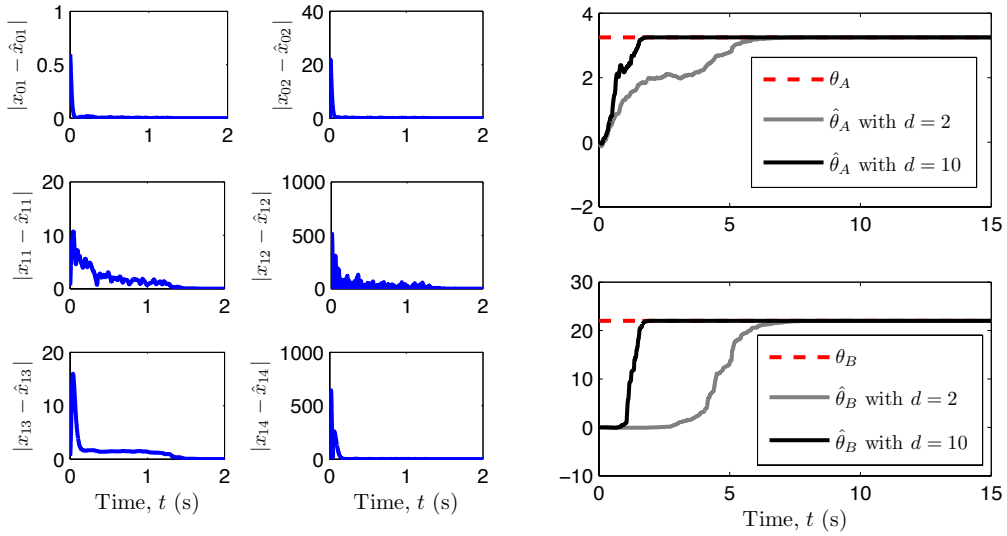
$$|(\tilde{x}(t), \tilde{p}(t))| \leq \beta_d\Big(|(\tilde{x}(0), \tilde{p}(0))|, t \Big) \quad \forall t \geq 0, \tag{5.8}$$

where $\tilde{x} := x - \hat{x}$ and $\tilde{p} := p^\star - \hat{p}$. □

Remark 4. *The design parameter d is chosen such that $d \geq d^\star$ to obtain (5.8). An estimate of $d^\star$, which we believe is subject to some conservatism is provided in the proof of Theorem 6 (see (B.28) in the Appendix). Although the expression of $d^\star$ depends on $|p^\star|$, it is possible to estimate it by taking the maximum over the set Θ, where $p^\star$ is known to belong to according to Assumption 11.*

5.4 Simulations

We have performed simulations for the single cortical column model described in Section 5.1.1 with the following initial conditions: $x(0) = (0.6, 1, 0.6, 1, 0.6, 1)$, $\hat{x}(0) = 0$, $p^\star = (\theta_A, \theta_B) = (3.25, 22)$, $\hat{p}(0) = 0$, $P(0) = \mathbb{I}_2$, $\Upsilon(0) = 0$. The other parameters take the value given in [74] and the input is a Gaussian noise of mean 100 and variance 30^2. Figures 5.2(a)-5.2(b) respectively show the state and parameter estimation error when the design parameter d is equal to 2 and 10. They confirm the convergence properties of the algorithm and show that larger d speeds up the rate of convergence.



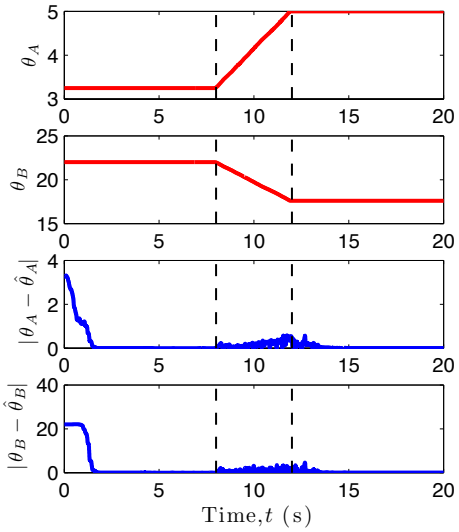
(a) State estimation error for $d = 10$.

(b) Parameter estimates for $d = 2$ and 10.

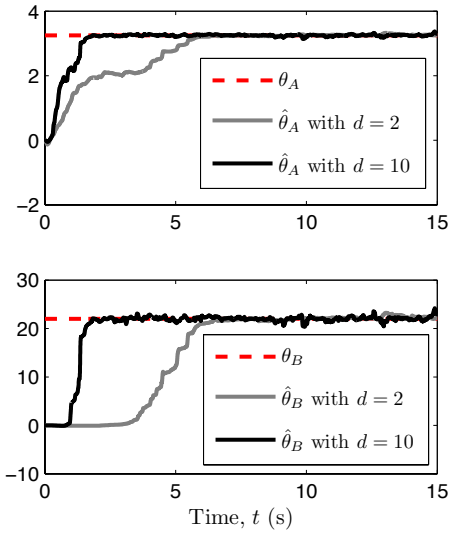
As previously mentioned, Assumptions 11 and 13 (i.e parameters are constant and noise-free measurement) are limiting assumptions in practice. In Figure 5.2(c), we simulate a change in parameters which might correspond to a change in cortical activity. We see that the estimated parameters still converge to the true values after the transition has occurred. We also consider the scenario where the measured output (EEG) is noisy by introducing a random noise that is drawn from a Gaussian distribution of mean 0 and variance 0.4^2, which is approximately 20% of y at steady state. Figure 5.2(d) shows that despite the presence of measurement noise, the adaptive observer still provides estimates that are close to the true values. In fact, the neighbourhood in which the parameter error converges to is smaller with small d. This illustrates the tradeoff between fast convergence rate and robustness towards measurement noise.

5.5 Summary

We have presented an algorithm to provide simultaneous estimates of the states and parameters of a class neural mass models using the electroencephalogram (EEG) as the measured output. This adaptive observer is based on the work in [46], where the class of systems considered are different in the linear part. Consequently, modifications were made in the adaptive observer and we provided a technical proof for the asymptotic



(c) Parameter estimation error with varying parameters for $d = 10$.



(d) Parameter estimates with noisy measurement for $d = 2$ and 10.

convergence of parameters and states. Simulations are provided for a single cortical column model by [74] which confirm our results. As with the state observer designed in Chapter 4, the adaptive observer synthesised in this chapter is applicable to *any* system in the class considered (5.1) and not limited to the neural mass models in Chapter 2.

In reality, the parameters of the system varies as the brain transitions from one event to another, e.g. non-seizure to seizure activity. The assumption we made in our design (Assumption 11) is restrictive in that the parameters are assumed to be constant. However, we expect our results to hold provided that the parameters are slowly-varying, see Chapter 9.6 in [86] for a discussion on slowly varying systems.

We also employ the constant parameter assumption in the next chapter, where we provide an alternative estimation method for states and parameters using the supervisory architecture.

Chapter 6

Supervisory observer

In this chapter, we present a parameter and state estimation algorithm for *general* nonlinear systems, which we call a supervisory observer. We adopt an architecture known as the supervisory framework used in [95, Chapter 6] that is used for control (see [64; 66; 110; 152] for linear systems and [21] for nonlinear systems). In these works, the objective is to steer the plant's state to the origin and no guarantee is provided on the parameter and state estimates. Since our objective is to estimate and not to control, we adopt the supervisory setup without the multi-controller and adapt the available results, so that it is applicable for estimation.

The plant's parameters are assumed to be *constant* and they belong to a known, compact parameter set. We sample the parameter set to form a finite set of nominal parameter values. The distribution of these sampled parameters can be done based upon prior knowledge of the plant's unknown parameters. An observer with a robustness property is designed for each of the parameter values, which forms a bank of observers known as the *multi-observer*. The *supervisory* unit chooses an observer from the *multi-observer* based on a criterion using a logical decision rule that generates a piecewise constant signal that is the index of the chosen observer at every instant of time. We use the scale-independent hysteresis switching logic introduced in [64]. The selected observer at each time provides us with a state estimate and its corresponding parameter estimate. In this chapter, we assume that the solution of the plant's state is uniformly bounded and all the output estimation errors of the multi-observer satisfy a persistency of excitation (PE) property. The parameter and state estimation error are guaranteed to converge to the origin with some margin of error, provided that the sampling of the parameter set is sufficiently large such that the distance from the plant's parameter to the sampled parameter set is small enough.

We believe that the advantages of supervisory control discussed in [65] translate well to the problem of state and parameter estimation. Firstly, it is not necessary to construct an adaptive observer, which is a challenging problem for nonlinear systems. In fact, very few *provable* designs exist, especially for nonlinearly parameterised systems (see [46] and the references there-in). Secondly, the supervisory framework has the advantage of modularity. Each of the components: the switching logic, the monitoring signals and the multi-observer can be designed independently to satisfy the respective properties required to meet our objective. This allows for the usage of readily available *state* observers for the additional purpose of parameter estimation. In Section 6.4, we design Luenberger observers for linear systems and a robust form of circle criterion observers [34] (Chapter 4.2) for a class of nonlinear systems. Finally, in Section 6.5, we show that this setup can be applied to the neural mass models in Chapter 2.

6.1 Problem formulation

We consider the following class of nonlinear systems

$$\begin{aligned} \dot{x} &= f(x, p^\star, u) \\ y &= h(x, p^\star), \end{aligned} \tag{6.1}$$

where the state vector is $x \in \mathbb{R}^{n_x}$, the output is $y \in \mathbb{R}^{n_y}$, the input is $u \in \mathbb{R}^{n_u}$ and the *unknown* parameter vector $p^\star \in \Theta \subset \mathbb{R}^{n_p}$ is constant, where Θ is a *known* compact set. For any initial condition $x(0)$ and input $u \in \mathcal{M}_{\Delta_u}$, where $\Delta_u > 0$, system (6.1) admits a unique solution $x(\cdot, x(0), u_{[0,t]}, p^\star)$ that is defined for all positive time. The short hand notation $x(\cdot)$ will often be used instead when there is no ambiguity from the context. The following additional assumptions are made on the plant.

Assumption 15. *The function $h : \mathbb{R}^{n_x} \times \mathbb{R}^{n_p} \to \mathbb{R}^{n_y}$ is continuously differentiable.* □

Assumption 16. *The solutions of system (6.1) are uniformly bounded, i.e for all Δ_x, $\Delta_u \geq 0$, there exists a constant $K_x = K_x(\Delta_x, \Delta_u) > 0$ such that for all $(x(0), u) \in B_{\Delta_x} \times \mathcal{M}_{\Delta_u}$*

$$|x(t)| \leq K_x, \qquad \forall t \geq 0. \tag{6.2}$$

□

At this stage, it is important to note that the neural mass models discussed in Chapter 2 satisfy this assumption. We will show how to apply the general results to the model by Jansen and Rit later in Section 6.5.

Remark 5. *Contrary to the problem of supervisory control in [151] and [152], we do not require the system (6.1) to be stabilisable. Since our purpose is to estimate, we only require the solutions of system (6.1) to be uniformly bounded.*

The main objective of this paper is to estimate the parameters $p^\star$ and the states x of system (6.1) when only the input u and the output y of system (6.1) is measured, using the supervisory observer which is described in the next section.

6.2 Supervisory observer

Based on the supervisory framework used for control in [152], the proposed methodology consists of two basic units: a bank of observers (multi-observer) which generates state estimates and the supervisor (monitoring signals and switching logic) which chooses one observer as shown in Figure 6.1. The estimated parameters and states are derived from the choice the supervisor makes.

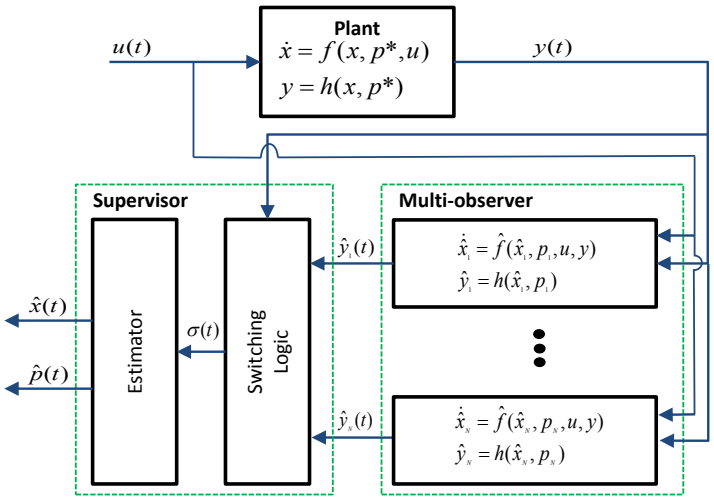


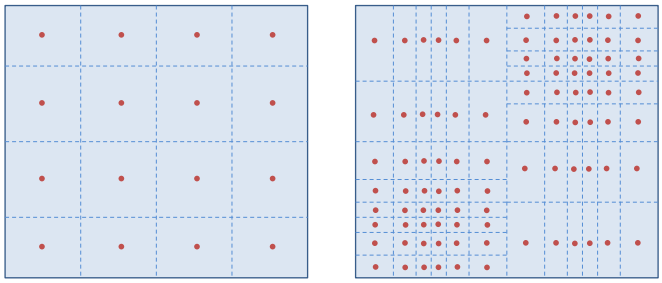
Figure 6.1: Supervisory observer.

6.2.1 Sampling of the parameter set Θ

We choose N distinct parameter values from the known, compact parameter set Θ. Thus, they form a sampled parameter set denoted $\Theta_o := \{p_1, \ldots, p_N\} \subset \Theta$. We denote the induced parameter error set as $\tilde{\Theta} := \{p_i - p^\star : p^\star \in \Theta, p_i \in \Theta_o, i \in \{1, \ldots, N\}\}$.

The distribution of the sampled parameters can be logarithmic as shown in Figure 6.2, where more p_i's are chosen in regions where the plant's parameter $p^\star$ is expected to

be. On the same vein, the sampling may also be done according to a known probability distribution of the plant's parameters, where more samples are taken in regions of the parameter set with the higher probability. In cases where no prior information on the plant's parameter $p^\star$ is known, uniform sampling may be implemented, where each p_i's are placed equidistance apart as shown in Figure 6.2.



(a) Uniform (b) Logarithmic

Figure 6.2: Sampling of the parameter set $\Theta \subset \mathbb{R}^2$. Dots represent the selected parameter p_i.

6.2.2 Multi-observer

For each parameter p_i, $i \in \{1, \ldots, N\}$, a state observer is designed

$$\begin{aligned} \dot{\hat{x}}_i &= \hat{f}(\hat{x}_i, p_i, u, y) \\ \hat{y}_i &= h(\hat{x}_i, p_i), \end{aligned} \tag{6.3}$$

where $\hat{x}_i \in \mathbb{R}^{n_x}$ is the state estimate. The solutions of (6.3) are assumed to be unique and defined for all time for all initial conditions $\hat{x}_i(0)$, any input u, plant's output y and parameter $p_i \in \Theta_o$. Denoting the state estimation error as $\tilde{x}_i := \hat{x}_i - x$, the output error as $\tilde{y}_i := \hat{y}_i - y$ and the parameter error as $\tilde{p}_i := p_i - p^\star$, we obtain the following state estimation error systems

$$\begin{aligned} \dot{\tilde{x}}_i &= \hat{f}(\tilde{x}_i + x, \tilde{p}_i + p^\star, u, y) - f(x, p^\star, u) =: F_i(\tilde{x}_i, \tilde{p}_i, p^\star, u, x) \\ \tilde{y}_i &= h(\tilde{x}_i + x, \tilde{p}_i + p^\star) - h(x, p^\star) =: H(\tilde{x}_i, x, \tilde{p}_i, p^\star). \end{aligned} \tag{6.4}$$

We assume the observers (6.3) are designed such that the following property holds.

Assumption 17. *Consider the state estimation error system (6.4) for*

$i \in \{1, \ldots, N\}$. There exists a continuously differentiable function $V_i : \mathbb{R}^{n_x} \to \mathbb{R}_{\geq 0}$, a continuous positive function $\tilde{\gamma} : \tilde{\Theta} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}_{\geq 0}$ where $\tilde{\gamma}(0, z, \bar{z}) = 0$ for all $z \in \mathbb{R}^{n_x}$, $\bar{z} \in \mathbb{R}^{n_u}$, and constants a_1, a_2, $\lambda_0 > 0$, $\mu_\nu \geq 1$ such that the following holds for all $\tilde{p}_i \in \tilde{\Theta}$, for all $u \in \mathbb{R}^{n_u}$, $\tilde{x}_i \in \mathbb{R}^{n_x}$, $x \in \mathbb{R}^{n_x}$

$$a_1 |\tilde{x}_i|^2 \leq V_i(\tilde{x}_i) \leq a_2 |\tilde{x}_i|^2, \tag{6.5}$$

$$\frac{\partial V_i}{\partial \tilde{x}_i} F_i(\tilde{x}_i, \tilde{p}_i, p^\star, u, x) \leq -\lambda_0 V_i(\tilde{x}_i) + \tilde{\gamma}(\tilde{p}_i, x, u). \tag{6.6}$$

□

Assumption 17 implies that the origin of the state estimation error system (6.4) is globally exponentially stable when $p_i = p^\star$, i.e. when we know the value of the plant's parameters $p^\star$. When this is not the case, condition (6.6) needs to be satisfied, which is a robustness property of system (6.4) with respect to $\tilde{p}_i$. It is noted that Assumption 17 is satisfied when system (6.4) is input-to-state *exponentially* stable [58, Equation (5)] with respect to $\tilde{p}_i$. In Section 6.4, we will show that Luenberger observers and circle criterion observers satisfy Assumption 17.

Following the terminology used in [152], we call the bank of observers in (6.3) a multi-observer.

Remark 6. *If the function $\tilde{\gamma}$ in (6.6) is dependant only on $\tilde{p}_i$ and u, Assumption 16 can be relaxed in the sense that the solutions of (6.1) do not have to be uniformly bounded. In this case, for all $u \in \mathcal{M}_{\Delta_u}$, the system (6.4) is locally input-to-state stable (ISS) [85, Theorem 5.2] with respect to $\tilde{p}_i$.*

Remark 7. *The Lyapunov-based conditions (6.5)-(6.6) stated in Assumption 17 differ from the conditions stated in [152, Equations (10a) and (10b)] and [151, Theorem 4.3 (iii)] because our objective is the estimation of the parameters and states of (6.1), whereas the available results in the literature consider the problem of stabilisation of an equilibrium point without parameter estimation. Moreover, in the literature, the state of the stabilised plant converges to zero, whereas in our case, the state of the plant does not necessarily converge to zero.*

6.2.3 Monitoring signal

As used in [152, Equation 6], the monitoring signal associated with each observer is the exponentially weighted $\mathcal{L}_2$ norm [71] of the output error $\tilde{y}_i$ defined as

$$\mu_i(t) = \int_0^t \exp(-\lambda(t-s))|\tilde{y}_i(s)|^2 ds + c_\mu, \qquad \forall t \geq 0, \tag{6.7}$$

where constants λ and $c_\mu > 0$ are design parameters. The signal μ_i can be implemented with c_μ added to the output of a linear filter

$$\dot{\pi}_i = -\lambda \pi_i + |\tilde{y}_i|^2, \tag{6.8}$$

with $\pi_i(0) = 0$. We assume that the output error of each of the observers $\tilde{y}_i$ satisfies the following property.

Assumption 18. *Consider the state error system (6.4) for all $i \in \{1, \ldots, N\}$. For all $\Delta_{\tilde{x}}$, Δ_x, $\Delta_u > 0$, there exist a constant $T_f = T_f(\Delta_{\tilde{x}}, \Delta_x, \Delta_u) > 0$ and a class $\mathcal{K}$ function $\alpha_{\tilde{y}}(\Delta_{\tilde{x}}, \Delta_x, \Delta_u)$ such that for all $\tilde{x}_i(0) \in B_{\Delta_{\tilde{x}}}$, $x(0) \in B_{\Delta_x}$, for some $u \in \mathcal{M}_{\Delta_u}$, and for all $\tilde{p}_i \in \tilde{\Theta}$, the corresponding solution to (6.4) satisfies*

$$\int_{t-T_f}^t |\tilde{y}_i(\tau)|^2 d\tau \geq \alpha_{\tilde{y}}(|\tilde{p}_i|), \qquad \forall t \geq T_f. \tag{6.9}$$

□

The inequality (6.9) is known as a *persistency of excitation (PE) condition* that appears in identification and adaptive literature [71]. It differs from the classical PE definition[1] in that we consider here a family of systems parameterised with $\tilde{p}_i$, where the lower bound (excitation level) in (6.9) depends on $\tilde{p}_i$. In particular, if $\tilde{p}_i = 0$, we do not necessarily have PE in the classical sense. In addition, the excitation level grows with the norm of the parameter error $\tilde{p}_i$. Hence, the integral term in (6.9) provides some quantitative information about the parameter estimation error. Assumption 18 holds when the output errors $\tilde{y}_i$, for $i \in \{1, \ldots, N\}$ satisfy the classical PE condition (6.10) according to the proposition below.

[1] A vector signal $\phi : \mathbb{R}_{\geq 0} \to \mathbb{R}^n$ is said to be *persistently exciting* (PE) with *excitation level* $\alpha > 0$ if there exists a constant $T > 0$ such that for all $t \geq T$

$$\int_{t-T}^t \phi(s)\phi(s)^T \, \mathrm{ds} \geq \alpha \mathbb{I}. \tag{6.10}$$

Proposition 2. *Consider the state error system (6.4) for $i \in \{1, \dots, N\}$. Suppose that the following holds.*

1. *Assumption 15 is satisfied.*
2. *The function $f : \mathbb{R}^n \times \mathbb{R}^{n_p} \times \mathbb{R}^{n_u} \to \mathbb{R}^n$ is continuously differentiable.*
3. *For all $\Delta_{\tilde{x}}$, Δ_x, $\Delta_u > 0$ and for all $\tilde{p}_i \neq 0$ with $\tilde{p}_i \in \tilde{\Theta}$, there exist constants $T_f = T_f(\Delta_{\tilde{x}}, \Delta_x, \Delta_u, \tilde{p}_i)$, $\bar{\alpha}_i = \bar{\alpha}_i(\Delta_{\tilde{x}}, \Delta_x, \Delta_u, \tilde{p}_i) > 0$ such that for all $\tilde{x}_i(0) \in B_{\Delta_{\tilde{x}}}$, $x(0) \in B_{\Delta_x}$ and for some $u \in \mathcal{M}_{\Delta_u}$, the corresponding solution to (6.4) satisfies the following*
$$\int_{t-T_f}^{t} |\tilde{y}_i(\tau)|^2 d\tau \geq \bar{\alpha}_i, \quad \forall t \geq T_f. \tag{6.11}$$

Then (6.9) holds. □

Proposition 2 indicates that tools to verify the classical PE condition (6.11) (see [112, Chapter 6] for linear systems and [118] for nonlinear systems) can be applied to ensure that Assumption 18 is satisfied.

6.2.4 Switching logic

The switching logic $\sigma : [0, \infty) \to \{1, \dots, N\}$ produces a piecewise constant signal which chooses an observer from the bank of N observers at every instant of time. We implement the scale-independent hysteresis switching logic [64]

$$\sigma(t) := \begin{cases} \underset{i \in \{1,\dots,N\}}{\arg\min}\, \mu_i(t), & \exists j \in \{1, \dots, N\} \text{ such that } (1+h)\mu_j(t) \leq \mu_{\sigma(t_s^-)}(t) \\ \sigma(t_s^-), & \text{otherwise,} \end{cases} \tag{6.12}$$

where t_s denotes the latest switching time of σ before t. On the occasion where $\arg\min \mu_i$ is non-unique, the switching logic $\sigma(t)$ selects the smallest index $i \in \{1, \dots, N\}$ which satisfies $\arg\min \mu_i$. The hysteresis constant $h > 0$ is a design parameter.

The switching signal σ implemented with the monitoring signals defined in (6.7) has an average dwell-time (see [95, Section 3.2.2]), i.e. there exist τ_a, $N_0 > 0$ such that the number of discontinuities (number of switchings) on an arbitrary time interval $[t_0, t)$ denoted by $N_\sigma(t, t_0)$ satisfies the following

$$N_\sigma(t, t_0) \leq N_0 + \frac{t - t_0}{\tau_a}. \tag{6.13}$$

This means that there is a finite number of switches in every finite interval, thereby no Zeno behaviour can occur. The following lemma is shown along the same lines as [63; 151].

Lemma 1. *Consider the plant* (6.1)*, the multi-observer* (6.3)*, the monitoring signal* (6.7) *and the switching logic* (6.12) *under Assumptions 15-18. The number of switchings of the switching logic* σ *satisfies* (6.13) *with*

$$\tau_a := \frac{\ln(1+h)}{\lambda N}, \quad N_0 \geq 1, \tag{6.14}$$

where $h > 0$ *is the hysteresis constant from* (6.12)*,* $\lambda > 0$ *is from the monitoring signal* (6.7) *and* $N \geq 1$ *is the number of observers in the multi-observer* (6.3)*.* □

6.2.5 Parameter and state estimates

Based on the signal generated by the switching logic σ from (6.12), the estimated parameters $\hat{p} : [0, \infty) \to \{p_1, \ldots, p_N\}$ and estimated states $\hat{x} : [0, \infty) \to \mathbb{R}^n$ are

$$\hat{p}(t) := p_{\sigma(t)}, \tag{6.15}$$

$$\hat{x}(t) := \hat{x}_{\sigma(t)}(t), \qquad \forall t \geq 0. \tag{6.16}$$

The state estimate $\hat{x}$ is picked from the solutions of a family of observer systems (6.3), for $i \in \{1, \ldots, N\}$. The chosen state estimate $\hat{x}(t) = \hat{x}_{\sigma(t)}(t)$ as defined in (6.16) is discontinuous in general, which is caused by the switching between different state estimates that are in general different at the switching instant t_s.

Remark 8. *Choosing a solution from a family of systems*

$$\begin{cases} \dot{\hat{x}}_1 &= \hat{f}(\hat{x}_1, p_1, u, y) \\ \vdots & \vdots \\ \dot{\hat{x}}_N &= \hat{f}(\hat{x}_N, p_N, u, y), \end{cases} \tag{6.17}$$

in the way the estimated state $\hat{x}_{\sigma(t)}$ *is chosen differs from a switched system with state jumps, in that we switch among a family of observer systems* (6.3) *whose solutions do not get affected by the switching. To be precise, let* $\hat{x}_i\Big(t; \hat{x}_i(0)\Big)$*, for* $i \in \{1, \ldots, N\}$ *denote the solution of the observer system* (6.3) *for any* $\hat{x}_i(0) \in B_{\Delta_x}$ *and* $u \in \mathcal{M}_{\Delta_u}$*, where we have omitted the dependance on* u*,* y *and* p_i*. Consider the switching logic* σ *with a switching sequence* $\{t_1, t_2\}$*, where* $t_2 > t_1 > 0$ *and* $\sigma(0) = 1$*,* $\sigma(t_1) = 2$ *and*

$\sigma(t_2) = 1$. *Our estimated state* $\hat{x}_{\sigma(t)}(t)$ *for all* $t \geq 0$ *is*

$$\hat{x}_{\sigma(t)}(t) = \begin{cases} \hat{x}_1\Big(t; \hat{x}_1(0)\Big), & t \in [0, t_1) \\ \hat{x}_2\Big(t; \hat{x}_2(0)\Big), & t \in [t_1, t_2) \\ \hat{x}_1\Big(t; \hat{x}_1(0)\Big), & t \in [t_2, \infty). \end{cases} \tag{6.18}$$

This is different from a switched system with state jumps where its solution for all $t \geq 0$ *is*

$$\hat{x}_{\sigma(t)}(t) = \begin{cases} \hat{x}_1\Big(t; \hat{x}_1(0)\Big), & t \in [0, t_1) \\ \hat{x}_2\Big(t; \varphi\big(\hat{x}_1(t_1)\big)\Big), & t \in [t_1, t_2) \\ \hat{x}_1\Big(t; \varphi\big(\hat{x}_2(t_2)\big)\Big), & t \in [t_2, \infty), \end{cases} \tag{6.19}$$

where φ *is some jump map.*

This distinction is crucial because provided that σ *is a piecewise continuous signal,* $\hat{x}_{\sigma(t)}$ *in* (6.18) *retains the behaviour of the individual systems* (6.3)*, whereas the solutions of a switched system with state jumps in* (6.19) *may exhibit a different behaviour as discussed in Part II of [95].* □

6.3 Main result

Before stating our main result, we first provide some intuition behind our result. Guidelines for the user are then provided to assist in implementing the supervisory observer.

Recall that N distinct parameters $p_i \in \Theta_o$ for $i \in \{1, \ldots, N\}$ are chosen from the known parameter set Θ and they form the sampled parameter set $\Theta_o \subset \Theta$. Let the point-to-set distance of the plant's parameter $p^\star$ to the set of sampled parameters Θ_o be

$$d(p^\star, \Theta_o) := \min_{p_i \in \Theta_o \text{ for } i \in \{1,\ldots,N\}} |p^\star - p_i|. \tag{6.20}$$

We are interested in the scenario where the selected parameters p_i's are distributed in Θ in a manner that $d(p^\star, \Theta_o)$ decreases as we increase the number of observers N. Most notably, as $N \to \infty$, $d(p^\star, \Theta_o) \to 0$, which forms the main intuition behind Theorem 7 whose proof can be found in Appendix B.10.

Theorem 7. *Consider the plant* (6.1)*, the multi-observer* (6.3)*, the monitoring signals* (6.7)*, the switching logic* (6.12)*, the parameter estimate* (6.15) *and the state estimate* (6.16)*. Suppose Assumptions 15-18 hold. For any* $\Delta_{\tilde{x}}$, Δ_x, $\Delta_u > 0$, *for all* $\nu_{\tilde{x}}$, $\nu_{\tilde{p}} > 0$, *for all* λ, $c_\mu > 0$ *(as in* (6.7)*), there exist constants* $\bar{K}_{\tilde{x}}$, $\delta^\star$ *and* $h^\star > 0$, *such that for*

all $h \in (0, h^\star]$*, for any* Θ_o *with* $d(p^\star, \Theta_o) \leq \delta^\star$*, there exists a constant* $T > 0$*, such that for all* $(x(0), \tilde{x}_{\sigma(0)}(0)) \in B_{\Delta_x} \times B_{\Delta_{\tilde{x}}}$ *and for any* $u \in \mathcal{M}_{\Delta_u}$ *that satisfies Assumption 18, the following holds*

$$|\tilde{p}_{\sigma(t)}| \leq \nu_{\tilde{p}},\ \forall t \geq T, \qquad \limsup_{t\to\infty} |\tilde{x}_{\sigma(t)}(t)| \leq \nu_{\tilde{x}}, \qquad |\tilde{x}_{\sigma(t)}(t)| \leq \bar{K}_{\tilde{x}},\ \forall t \geq 0. \tag{6.21}$$

where $\tilde{p}_{\sigma(t)} := \hat{p}(t) - p^\star$ *is the parameter estimation error and* $\tilde{x}_{\sigma(t)} := \hat{x}(t) - x(t)$ *is the state estimation error.* □

Theorem 7 shows that the parameter and state estimates converge to the true values with desired accuracy $\nu_{\tilde{p}}$ and $\nu_{\tilde{x}}$ respectively, provided that the parameter set Θ is sampled sufficiently, i.e. the number of observers N is sufficiently large. Moreover, the state estimation error $\tilde{x}_\sigma$ is guaranteed to be bounded.

We provide some guidelines for the user to implement the algorithm below.

1. The system (6.1) is first verified to satisfy Assumptions 15-16.

2. *Selection of parameter values* p_i: a finite number of parameter values are selected from the parameter set Θ. Our result hinges upon taking a sufficiently large sampling of the parameter set in the way that $d(p^\star, \Theta_o)$ becomes small enough such that the conclusions of Theorem 7 hold ($d(p^\star, \Theta_o) \leq d^\star$ in Theorem 7). This may require high computing resources in practice, however the sampling of the parameter set can be done in efficient ways as discussed earlier in Section 6.2.1.

3. *Design of the multi-observer* : For each of the sampled parameter values, we design a state observer (6.3) such that Assumption 17 holds. The synthesis of a nonlinear state observer is a difficult problem in general. Nevertheless, various designs are available in the literature for specific classes of systems [23]. Examples are provided in Section 6.4. Typically, an observer is first designed under the assumption that the plant's parameter is known such that the equilibrium of the state estimation error system is globally exponentially stable. We then either verify or design each observer such that it satisfies the robustness property (6.6).

4. *Satisfaction of the PE condition* (6.9) : Assumption 18 may be satisfied with the aid of Proposition 2. The reader is referred to [112, Chapter 6] and [118] for properties of PE signals of linear and nonlinear systems respectively to obtain a priori checkable conditions. Otherwise, Assumption 18 may be checked in simulation a posteriori since we have access to $\tilde{y}_i$ online as illustrated by an example from neuroscience in Section 6.5.

5. *Choice of design parameters h, λ and c_μ* : The parameters λ and c_μ relates directly to the properties of the switching signal σ. In Lemma 1, it is shown that the average dwell time of the switching logic σ is $\tau_a := \frac{\ln(1+h)}{\lambda N}$. Hence, increasing λ and choosing a small enough hysteresis constant h (from (6.12)) speeds up the convergence rate of the parameters.

Consequently, the estimated parameter converges to a ball of a given margin $\nu_{\tilde{p}}$ centered at the true parameter $p^\star$ in finite time and the state estimation error also converges to a ball of a given margin $\nu_{\tilde{x}}$ as $t \to \infty$. Additionally, the state estimation error is bounded.

6.4 Applications

In this section, we investigate two case studies: (i) linear systems, for which Luenberger observers are designed and (ii) a class of nonlinear systems, for which robust circle criterion observers (first introduced in [17] and a robust version introduced in [34]) are synthesised.

6.4.1 Linear systems

We consider a linear plant

$$\begin{aligned} \dot{x} &= A(p^\star)x + B(p^\star)u \\ y &= C(p^\star)x, \end{aligned} \tag{6.22}$$

where $x \in \mathbb{R}^{n_x}$, $u \in \mathbb{R}^{n_u}$, $p^\star \in \Theta \subset \mathbb{R}^{n_p}$ and $A(p)$ is assumed to be Hurwitz, for all $p \in \Theta$. We notice that Assumption 15-16 hold. Note that, although system (6.22) is linear, it is nonlinearly parameterised. We also assume the following.

Assumption 19. *The matrices $A(p)$, $B(p)$ are continuous in $p \in \Theta$.*

Each observer in the multi-observer is designed as follows for $i \in \{1, \dots, N\}$

$$\begin{aligned} \dot{\hat{x}}_i &= A(p_i)\hat{x}_i + B(p_i)u + L_i(p_i)(C(p_i)\hat{x}_i - C(p^\star)x) \\ \hat{y}_i &= C(p_i)\hat{x}_i, \end{aligned} \tag{6.23}$$

where $L_i(p_i)$ is such that $A(p_i) + L_i(p_i)C(p_i)$ is Hurwitz (this is always possible since $A(p_i)$ is Hurwitz, for all $p_i \in \Theta_o$, where $i \in \{1, \dots, N\}$). Denoting the state estimation

error as $\tilde{x}_i := \hat{x}_i - x$ and the parameter error as $\tilde{p}_i := p_i - p^\star$ as in Section 6.2, we obtain the following state estimation error system

$$\begin{aligned}\dot{\tilde{x}}_i &= \Big(A(p_i) + L_i(p_i)C(p_i)\Big)\hat{x}_i - \Big(A(p^\star) + L_i(p_i)C(p^\star)\Big)x + \Big(B(p_i) - B(p^\star)\Big)u \\ &= \Big(A(p_i) + L_i(p_i)C(p_i)\Big)\tilde{x}_i + \Big(\tilde{A}(p_i, p^\star) + L_i(p_i)\tilde{C}(p_i, p^\star)\Big)x \\ &\quad + \tilde{B}(p_i, p^\star)u, \end{aligned} \tag{6.24}$$

where we denote $\tilde{A}(p_i, p^\star) := A(p_i) - A(p^\star)$, $\tilde{B}(p_i, p^\star) := B(p_i) - B(p^\star)$ and $\tilde{C}(p_i, p^\star) := C(p_i) - C(p^\star)$. Proposition 3 shows that Assumption 17 is satisfied.

Proposition 3. *Consider the linear system (6.22), the multi-observer (6.23) and suppose Assumption 19 is satisfied. If there exist $P_i = P_i^T > 0$ and scalar constants ν_i, $\mu_i > 0$ such that the following holds*

$$\begin{bmatrix} P_i\left(A(p_i) + L_i(p_i)C(p_i)\right) + \left(A(p_i) + L_i(p_i)C(p_i)\right)^T P_i + \nu_i \mathbb{I} & P_i \\ P_i & -\mu_i \mathbb{I} \end{bmatrix} \leq 0. \tag{6.25}$$

Then Assumption 17 is satisfied. □

As explained earlier, if the classical PE condition (6.11) is guaranteed, Assumption 18 is satisfied according to Proposition 2. There exist results in the literature (see Chapter 6 in [112]) which provide sufficient conditions to verify (6.11) for linear systems. They can be used, for instance, to design an input to the system (6.22) such that the inequality (6.11) is satisfied. Lastly, since all the conditions of Theorem 7 are satisfied, we derive the following result.

Proposition 4. *Consider the linear system (6.22), the multi-observer (6.23) for $i \in \{1, \ldots, N\}$, the monitoring signals (6.7), the switching logic (6.12). Suppose that Assumptions 18 and 19 are satisfied, then the conclusions of Theorem 7 hold.* □

Proposition 4 extends the results in [89] from single-input single-output (SISO) to multi-input multi-output (MIMO) systems.

6.4.2 A class of nonlinear systems

We consider the following class of nonlinear systems first studied in [16] for the design of circle criterion observers

$$\begin{aligned}\dot{x} &= A(p^\star)x + G(p^\star)\gamma(Hx) + B(p^\star)\sigma(u, y) \\ y &= C(p^\star)x, \end{aligned} \tag{6.26}$$

where $\gamma : \mathbb{R}^{n_p} \to \mathbb{R}^{n_\gamma}$ and $\sigma : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \to \mathbb{R}^{n_\sigma}$. We assume that Assumption 16 holds and also make the following assumptions.

Assumption 20. *The matrices $A(p)$, $G(p)$ and $B(p)$ are continuous in $p \in \Theta$.*

Assumption 21. *For any $k \in \{1, \dots, n_\gamma\}$, there exist constants a_{γ_k}, $b_{\gamma_k} > 0$ such that the following holds*

$$-\infty < a_{\gamma_k} \leq \frac{\partial \gamma_k(v_k)}{\partial v_k} \leq b_{\gamma_k} < \infty, \qquad \forall v_k \in \mathbb{R}, \tag{6.27}$$

where $\gamma = (\gamma_1, \dots, \gamma_{n_\gamma})$. □

By Assumption 21, γ includes nonlinearities that are globally Lipschitz, with $b_{\gamma_k} = -a_{\gamma_k}$. This assumption was used in [158] and is a special case of the assumption in [17, Equation (1)] and [34, Assumption 1] for differentiable nonlinearities γ. Moreover, since system (6.26) is assumed to be bounded, there is no loss in generality in assuming that γ is globally Lipschitz (see for example Section 2 in [46]).

For each $p_i \in \Theta_o$, $i \in \{1, \dots, N\}$, the following observer [17; 34; 158] is designed

$$\begin{aligned} \dot{\hat{x}}_i &= A(p_i)\hat{x}_i + G(p_i)\gamma(H\hat{x}_i + K_i(C\hat{x}_i - y)) + B(p_i)\sigma(u, y) + L_i(C(p_i)\hat{x}_i - y) \\ \hat{y}_i &= C(p_i)\hat{x}_i, \end{aligned} \tag{6.28}$$

where K_i and L_i are observer matrices designed using the ideas in [34], i.e. such that a quadratic function $V_i : \mathbb{R}^{n_x} \to \mathbb{R}_{\geq 0}$ satisfies the conditions stated in Assumption 17.

Proposition 5. *Consider system (6.26) and multi-observer (6.28), for all $(x(0), \tilde{x}_i(0)) \in B_{\Delta_x} \times \mathbb{R}^{n_x}$ and $u \in \mathcal{M}_{\Delta_u}$. Under Assumptions 16, 20 and 21, if there exist matrices $P_i = P_i^T > 0$, $M_i = diag(m_{i1}, \dots, m_{in_p}) > 0$ and scalar constants ν_i, $\mu_i > 0$ such that the following holds*

$$\begin{bmatrix} \mathcal{A}(P_i, L_i, \nu_i) & \mathcal{B}(P_i, M_i, K_i^T M_i) & P_i \\ \star & \mathcal{E}(M_i) & 0 \\ \star & \star & -\mu_i \mathbb{I} \end{bmatrix} \leq 0, \tag{6.29}$$

where $\mathcal{A}(P_i, L_i, \nu_i) = P_i\big(A + L_i(p_i)C(p_i)\big) + \big(A + L_i(p_i)C(p_i)\big)^T P_i + \nu_i \mathbb{I}$,
$\mathcal{B}(P_i, M_i, K_i^T M_i) = P_i G(p_i) + \big(H + K_i(p_i)C(p_i)\big)^T M_i$ and
$\mathcal{E}(M_i) = -2M_i diag\left(\frac{1}{b_{\gamma_1}}, \dots, \frac{1}{b_{\gamma_{n_p}}}\right)$, then Assumption 17 is satisfied. □

We assume that the PE condition stated in Assumption 18 is satisfied. PE properties for nonlinear systems are studied in [118] and may be used in conjunction with

Proposition 2 to obtain checkable PE conditions for some classes of systems. Finally, we obtain the following result since all the conditions of Theorem 7 hold.

Proposition 6. *Consider the nonlinear system (6.26), the multi-observer (6.23), the monitoring signals (6.7), the switching logic (6.12) and suppose the following hold.*

1. *Assumptions 16, 18, 20 and 21 hold.*
2. *Condition (6.29) is feasible.*

Then the conclusions of Theorem 7 hold. □

Proposition 6 extends the results of [45, Theorem 3], [158] and [46], where adaptive observers are designed, in the sense that we consider a more general class of nonlinearly parameterised nonlinear systems.

6.5 Illustrative example: A neural mass model by Jansen and Rit

In this section, we apply the results of Section 6.4.2 to estimate the parameters and the states of the neural mass model in [74] and described in Chapter 2.3.2. It describes the dynamics of a single column model by capturing the interactions between the pyramidal neurons, the excitatory and the inhibitory interneurons in a localised region of the cortex. The model has been shown to realistically reproduce various patterns, such as alpha rhythms, seen in the electroencephalogram, also known as EEG (the measured output). Moreover, the model may be used to capture more complex phenomena as shown in [155]. The objective in this section is to estimate the mean membrane potential of neural populations (states x) and the synaptic gains of each of the populations (parameters $p^\star$).

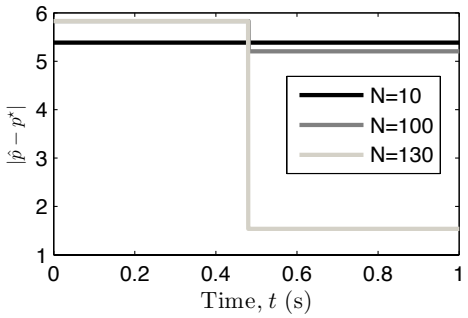
The model can be written in the form (6.26) as shown in Chapter 2.4.3. It also has uniformly bounded solutions for all initial conditions $x(0)$ and bounded input $u \in \mathcal{M}_{\Delta_u}$, because the matrix A is Hurwitz, the nonlinearity S in γ and σ is bounded. Therefore, the model satisfies Assumption 16. Furthermore, by the definition of the matrices A, G, B and the nonlinearity S, Assumptions 20 and 21 are satisfied.

We sample the parameter space non-uniformly by sampling p_2 more densely in the range of $20 - 30$ ($p_2^\star = 25$) and p_1 is distributed uniformly. The observers are designed as proposed in Section 6.4.2. Figure 6.3(d) shows that the classical PE condition (6.11) seems to be satisfied for when $N = 10$, 100 and 130. Hence, by Proposition 2, Assumption 18 may be satisfied. Therefore, the conclusions of Theorem 7 hold according to

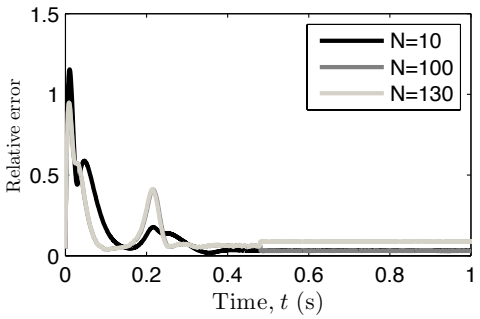
Proposition 6. We implement the supervisory setup with hysteresis constant $h = 0.5$, the design parameters from the monitoring signal (6.7) are $\lambda = 0.005$ and $c_\mu = 2$. Table 6.5 summarises the results for increasing N in such a way that $d(p^\star, \Theta_o)$ decreases. Figure 6.3 compares the performance of the algorithm for $N = 10$, 100 and 130.

N	10	50	100	120	130
$d(p^\star, \Theta_o)$	2.2361	1.1180	1.0244	1.0164	1.0035
$\limsup_{t\to\infty} \lvert\hat{p}(t) - p^\star\rvert$	5.385	5.385	5.204	5.233	1.537
$\limsup_{t\to\infty} \frac{\lvert\hat{x}(t)-x(t)\rvert}{\lvert\max_t x(t)-\min_t x(t)\rvert}$	0.033	0.033	0.031	0.031	0.0898

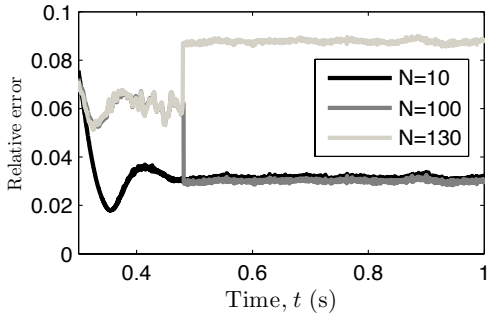
Table 6.1: Numerical results for increasing N values such that $d(p^\star, \Theta_o)$ decreases.



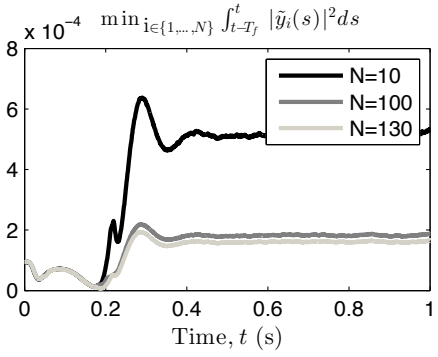
(a) Norm of parameter estimation error $|\hat{p}(t) - p^\star|$.



(b) Relative state estimation error $\frac{|\hat{x}(t)-x(t)|}{|\max_t x(t)-\min_t x(t)|}$.



(c) Relative state estimation error $\frac{|\hat{x}(t)-x(t)|}{|\max_t x(t)-\min_t x(t)|}$, for $0.3 \le t \le 1.0s$.



(d) Checking the PE condition in Assumption 18.

Figure 6.3: Simulation results.

From the simulation results, we see that the parameter estimation error converges to a neighbourhood of the origin whose size decreases with $d(p^\star, \Theta_o)$. The state estimation error also converges with some error, but it is interesting to see that the ultimate bound does not necessarily decrease with $d(p^\star, \Theta_o)$. This can be explained as follows. Although each observer satisfies Assumption 17, they do not have the same robustness properties with respect to the parameter error $\tilde{p}_i$. To be precise, $\bar{\gamma}_{\tilde{x}}$ of the individual state estimation error systems (B.35) are in general different. Thus, the decrease of $|\tilde{p}_i|$ induced by increasing N may be compensated by larger $\bar{\gamma}_{\tilde{x}}$. It will therefore be interesting to develop observers which minimise $\bar{\gamma}_{\tilde{x}}$ in future work.

The other neural mass models discussed in Chapter 2 also satisfy Assumptions 16, 20 and 21 as required and may satisfy Assumption 18 as discussed in the example. Hence, Proposition 5 may also be applicable to these models.

6.6 Summary

The main contribution of this chapter is the design of an algorithm for parameter and state estimation of *general* continuous-time nonlinear systems with uniformly bounded states and constant parameters that belong to a known, compact set. We use the scale-independent hysteresis switching logic to choose an observer from the multi-observer with robust properties. Provided that a PE condition is satisfied, we can guarantee the finite time practical convergence of the parameters and consequently, the asymptotic practical convergence of the states. We then investigate two case studies by applying our main result to linear systems and to a class of nonlinear systems. To illustrate our result, we consider the neural mass model by Jansen and Rit described in Chapter 2.3.2. Our proposed setup achieves asymptotic state estimation and *finite-time* parameter estimation within a margin of error that tends to 0 as the parameter set Θ is sampled sufficiently in an efficient way.

This chapter illustrates the potential of casting the problem of estimation of states and parameters in the hybrid systems setting. While hybrid tools have proved efficient for the control of continuous-time systems, few works have investigated it for estimation, see [62], [8] (parameter estimation only) and [80] (state estimation only). To the best of our knowledge, this is the first result that addresses both challenges from this perspective.

Chapter 7

Conclusions and future directions

ESTIMATION is a long-standing problem in control theory. This thesis presents state observers as well as state and parameter observers for nonlinear systems with a particular motivation of estimating neural activity from the EEG. For this purpose, we consider a class of neural mass models that includes models that replicate alpha rhythms in the cerebral cortex and also a model that captures seizure dynamics, to name a few.

In Part I: Chapters 3 and 4, we designed global exponential state observers for a class of nonlinear systems. We synthesised a globally exponential nonlinear estimator that is a posteriori robust towards disturbances and uncertainties in Chapter 3. This design is improved upon in Chapter 4 where we extended the circle criterion observer so that it is applicable to the class of systems we consider. Furthermore, we also took into account measurement disturbances and input uncertainties by redesigning the circle criterion observer such that their effects can be attenuated a priori by the user. The robustness property of these state observers is exploited later in Chapter 6.

Part II presents an adaptive observer in Chapter 5 and a supervisory observer in Chapter 6. The adaptive observer is designed for a class of nonlinear systems that is a subset of the class of neural mass models considered. We guarantee the asymptotic convergence of states and parameters to their true values. We present a parameter and state estimation methodology for *general* nonlinear systems in Chapter 6, which we call the supervisory observer. The supervisory observer uses readily available *state* observers such as the ones designed in Part I in a particular setup. Under certain conditions, we have finite-time practical convergence of parameters and asymptotic practical convergence of states to their true values.

The motivation behind this work stems from the exciting frontier of intersecting control theory and neuroscience for the prospect of developing new methodologies for

understanding neural phenomena and addressing neurological diseases such as epilepsy and Parkinson's disease. This thesis serves to provide mathematically proved estimation algorithms with tailored applications in the field of neuroscience, such as the study of alpha rhythm generation in the cortex and the anticipation of epileptic seizures. We stress that the observers designed in this thesis are also widely applicable to systems outside neuroscience that belong to the class of nonlinear systems considered. We summarise the limitations of this thesis and discuss future directions for research in neuroscience and nonlinear observer design in the sections that follow.

- This thesis considers only continuous-time systems. Evidently, the results presented are valid for discrete-time systems using discrete design tools. However, such designs and analyses are not included in this thesis.

- We deal only with systems represented by ordinary differential equations. These systems describe local behaviour such as neurons or cortical columns. An important extension of these models is the dynamics of neural activity on a spatially extended cortical sheet. This may be modelled by coupled cortical column models, where the coupling represents the strength of connectivity between different regions captured by each cortical column model. We considered interconnected cortical column models in Chapter 5, where we design an adaptive observer to estimate the states and parameters of each cortical column model under the assumption that the coupling between columns are known. Alternatively, the spatial dynamics of neural activity may be described by partial differential equations and they are known as neural field models (see [39]). Examples of such models include the models by Nunez [114] and Jirsa and Haken [77]. Observer designs for partial differential equations are available, see [92, Chapter 5] for an introduction, but is beyond the scope of this thesis.

- While this thesis provides mathematical proof of all designs and analysis of the observers presented, experimental verification on real EEG data remains to be done to validate the class of neural mass models considered.

7.1 Future directions in the applicability of model-based estimation in neuroscience

The implementation of the estimation algorithms on EEG data remains to verify the usability of this class of neural mass models for their respective purposes. Should this venture prove successful, then it forms the basis in achieving the long term goal of transforming theory to useful diagnosis, management and treatment tools in the clinical setting. In Part II of this thesis, we assumed that some of the parameters of the class of models considered (listed in Appendix C) are known and the model parameters that have been identified by their respective authors to be most relevant to the dynamics of the phenomenon of interest to be unknown. Our state and parameter observers presented in Part II are designed to estimate these unknown parameters. In practice, the known parameters may differ from one individual's brain to the other. Hence, these parameters would also need to be estimated. Considering these parameters to be unknown as well would result in a different class of systems compared to the ones considered in thesis. New observer designs would need to be devised to address them.

One of the many prospective outcomes of this work is the detection of epileptic seizures and eventually, the abatement of seizures. This is highly useful in the case of patients with highly localised seizures and who have pharmaco-resistant epilepsy or who are not candidates for resective surgery. Experimental works have been done in investigating the abatement of seizures by electrical stimulation (see for instance [49; 116]). However, these works often employ open-loop control, which is usually non-robust towards model uncertainty and measurement disturbances.

Closed-loop control addresses these issues. Works by [30; 31] use coupled neural mass models (a modified version of the model by Jansen and Rit [74] described in Chapter 2.3.2 is used in [31]) with the correlation between coupled model outputs as a measure of epileptic behaviour, $\rho : \mathbb{R}_{\geq 0} \to [0, 1]$. A proportional-integral (PI) controller is designed to regulate the coupling gain between two neural mass models via output feedback such that the correlation factor ρ is below an acceptable level where non-epileptic activity is exhibited.

A drawback of these studies is not exploiting the richness of behaviour this class of neural mass models is able to generate by varying their parameters. For instance, the model by Wendling et. al. [155] described in Chapter 2.3.1 is capable of generating EEG patterns that are usually seen during epileptic seizures via its parameters. Consider for

instance the model by Wendling et. al. written in the following form

$$\begin{aligned} \Sigma : \qquad \dot{x} &= f(x, u, p^*) \\ y &= h(x), \end{aligned} \tag{7.1}$$

where the state is $x \in \mathbb{R}^{n_x}$, the input is $u \in \mathbb{R}^{n_u}$, the parameter is $p^* \in \Theta \subset \mathbb{R}^{n_p}$, measurement/output is $y \in \mathbb{R}^{n_y}$ and n_x, n_u, n_p and n_y are positive integers. As mentioned in Chapter 2.3.1 and done in [155], the parameter set can be considered to be composed of non-seizure and seizure related parameters or more precisely, $\Theta := \Theta_S \cup \Theta_N$, where Θ_S is the seizure related parameter set and Θ_N is the non-seizure related parameter set.

A plausible problem formulation would be the design of the control input u such that the measurement y cease to exhibit seizure-like EEG patterns when $p^\star \in \Theta_s$. The statement 'the measurement y cease to exhibit seizure-like EEG patterns' may be interpreted as ensuring y tracks a particular set of trajectories that are known as non-epileptic behaviour. The control objective may be formulated as a problem of tracking or the stabilisation of the system with respect to a set. This setup is illustrated in Figure 7.1. Using the adaptive observer or supervisory observer already designed in this thesis (Chapters 5 and 6) to provide an estimate of the parameter $\hat{p}$, a switched control scheme outputs a control signal $u_c = 0$ when $\hat{p} \in \Theta_N$ and outputs a control signal u_c to be designed when $\hat{p} \in \Theta_S$. With this two-step approach of first designing the observer, then designing the control law, we need to ensure that the separation property holds in this context. Nonlinear control design tools are available to address such problems.

7.2 Future directions in nonlinear state and parameter observer design

In the direction of state and parameter estimation of nonlinear systems, many exciting further work will be pursued for the greater generality of these results. We address the extensions for the adaptive observer (Chapter 5) and the supervisory observer (Chapter 6) separately in the subsections that follow.

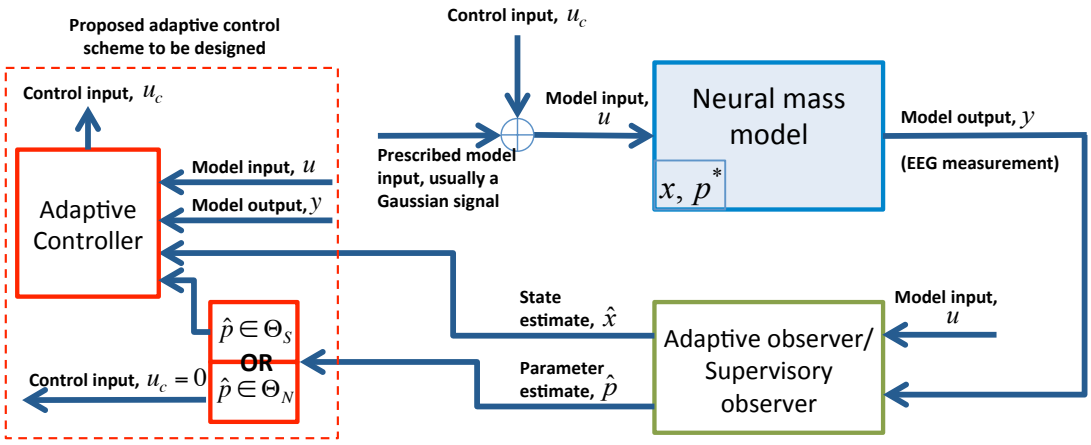


Figure 7.1: Proposed adaptive controller design for seizure abatement

7.2.1 Adaptive observer

We will work on extending the adaptive observer designed in Chapter 5 for the class of nonlinear systems (5.1) to the following more general class of nonlinear systems

$$\begin{aligned} \dot{x}_0 &= A_0 x_0 + \phi_0(y) p^\star \\ \dot{x}_1 &= A_1 x_1 + \phi_1(x_0, u) p^\star \\ \dot{x}_2 &= A_2 x_2 + \phi_2(x_0, x_1, u) p^\star \\ &\vdots \\ \dot{x}_n &= A_n x_n + \phi_n(x_0, x_1, \ldots, x_{n-1}, u) p^\star \\ y &= C_1 x_1 + C_2 x_2 + \cdots + C_n x_n, \end{aligned} \tag{7.2}$$

where the states are $x_i \in \mathbb{R}^{n_i}$ for $i \in \{0, \ldots, n\}$, the parameter vector is $p^\star \in \mathbb{R}^{n_p}$, the measurement is $y \in \mathbb{R}^{n_y}$ and the input is $u \in \mathbb{R}^{n_u}$. The nonlinear terms $\phi_0 : \mathbb{R} \to \mathbb{R}^{n_0} \times \mathbb{R}^{n_p}$ and $\phi_i : \mathbb{R}^{n_0} \times \cdots \times \mathbb{R}^{n_{i-1}} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_i} \times \mathbb{R}^{n_p}$, for $i \in \{1, \ldots, n\}$ are globally Lipschitz. The matrices $A_0, \ldots, A_n$ are Hurwitz.

This extension widens the class of systems considered by also including the model by Wendling et al [155] described in Chapter 2.3.1. The available work in [46] applies to a class of systems very similar to (7.2), but the issue arises in the linear part and the output that is not of the same form as (7.2). Hence, future work entails addressing

this problem.

Another aspect that needs to be addressed is the robustness of the adaptive observers with respect to disturbances and uncertainties. We may perform a posteriori analysis or re-design the observer such that the desired robustness properties are obtained. This is essential for its practical applicability in realistic scenarios.

7.2.2 Supervisory observer

There is scope for much future research in supervisory estimation. In the results presented in Chapter 6, we have limited ourselves to considering constant plant parameters $p^\star$. The conclusions of Theorem 7 may also be modified to for the case of slowly-varying parameters $p^\star$ using results for slowly-varying systems (see [86, Chapter 9.6]).

The robustness of the setup with respect to unmodelled dynamics and disturbances in measured output y and input u remains to be investigated. This robustness property is useful for the implementation of the supervisory observer in a practical setting and will also form the basis for investigating other forms of disturbances in measured output y, for example sampled and quantised measurements. This will have practical applications, for instance when y is acquired by a digital device, such as in the case of an implanted seizure control device or when y is transmitted over a limited capacity channel.

Further work will also be done in obtaining conditions where the Assumption 18 (PE) is guaranteed. Under certain conditions, an input u may be designed with the properties in [112, Chapter 6] for linear systems and [118] for nonlinear systems such that Assumption 18 holds.

One drawback of the supervisory observer is the potentially high computation needed for estimation as a result of the high number of parameter set sampling required to obtain desirable ultimate bounds of the state and parameter estimation errors. This is prohibitive when implemented with a high dimensional model, such as in the scale of the full brain, instead of a local region of the cerebral cortex. Furthermore, if the computations are to be executed by an implantable device, the high power consumption due to the large number of computations is undesirable. Nevertheless, if prior information about the plant's parameter is available (e.g. the probability distribution of the plant's parameter in the set), then computational complexity may be reduced by sampling efficiently according to the information available.

7.3 Concluding remarks: E tenebris in lucem

Control theoretic approaches to answering questions in neuroscience are still in its infancy. This thesis presents some methodologies for estimating neural activity from measurements such as the EEG. We hope that this endeavour will aid in better understanding of the human brain that eventually translates to useful medical solutions in anticipating and controlling seizures caused by neurological disorders.

Appendix A: Mathematical preliminaries

This appendix provides important tools used in the design of nonlinear observers and for analysing the stability of systems in this thesis. A variety of references were used to compile this appendix, the most important one is the textbook by H. Khalil [86].

Consider the following ordinary differential equation

$$\dot{x} = f(x, t), \tag{A.3}$$

where the vector $x(t) \in \mathbb{R}^{n_x}$ is the state for every $t \geq t_0 \geq 0$. The solution of (A.3) with initial condition $x(t_0)$ is denoted $x(t, x(t_0))$. The short hand notation $x(t)$ will be used instead where there is no ambiguity.

A model of a physical system such as a region of the human brain can only make sense if $x(t)$ exists, for all $t \geq t_0 \geq 0$. This system is said to be *forward complete* [96]. For this to be true and for $x(t)$ to be unique for all $t \geq t_0 \geq 0$, f is assumed to satisfy the following properties, for all initial conditions $x(t_0) \in B_{\Delta_x} \subset \mathbb{R}^{n_x}$:

- there exists $L > 0$ such that

$$|f(t, x) - f(t, y)| \leq L|x - y|, \qquad \forall (x, y, t) \in B_{\Delta_x} \times B_{\Delta_x} \times \mathbb{R}_{\geq 0}. \tag{A.4}$$

- $f(t, x)$ is piecewise continuous in t and continuous in x.

The first condition is known as the Lipschitz property of the function f with Lipschitz constant L.

An equilibrium point $x_e \in \mathbb{R}^n$ of (A.3) is obtained by solving $f(t, x_e) = 0$, for all

$t \geq t_0$. A non-zero equilibrium point can always be translated to the origin, which gives the following stability notions in terms of the origin $x_e = 0$ as the equilibrium point:

Definition 2 (Stability notions)**.** *The origin $x_e = 0$ of (A.3) is*

- stable *if, for each $\epsilon > 0$, there is a $\delta = \delta(\epsilon, t_0) > 0$ such that*

$$|x(t_0)| \leq \delta \implies |x(t)| < \epsilon, \qquad \forall t \geq t_0 \geq 0. \tag{A.5}$$

- uniformly stable *(US) if, for each $\epsilon > 0$, there is a $\delta = \delta(\epsilon) > 0$ independent of t_0, such that (A.5) is satisfied.*

- unstable *if it is not stable.*

- asymptotically stable *if it is stable and there is a positive constant $c = c(t_0)$ such that $x(t) \to 0$ as $t \to \infty$, for all $|x(t_0)| < c$.*

- uniformly asymptotically stable *(UAS) if it is US and there is a positive constant c independent of t_0, such that for each $\eta > 0$, there is a $T = T(\eta) > 0$ such that*

$$|x(t)| < \eta, \forall t \geq t_0 + T(\eta), \forall |x(t_0)| < c. \tag{A.6}$$

- globally uniformly asymptotically stable *(UGAS) if it is US, $\delta(\epsilon)$ satisfies $\lim_{t\to\infty} \delta(\epsilon) = \infty$ and for η, $c > 0$, there is a $T = T(\eta, c) > 0$ such that*

$$|x(t)| < \eta, \qquad \forall t \geq t_0 + T(\eta, c), \forall |x(t_0)| < c. \tag{A.7}$$

The $\epsilon - \delta$ definition of different stability notions can be made more transparent using comparison functions, i.e. class $\mathcal{K}$ functions and class $\mathcal{KL}$ functions in the following lemma. We state this for the definitions of US and UAS as well as a special case of UAS.

Lemma 2 (Lemma 4.5 and Definition 4.5 in [86]). *The equilibrium point $x_e = 0$ of (A.3) is*

- *US if and only if there is a class $\mathcal{K}$ function α and a positive constant c independent of t_0 such that*

$$|x(t)| \leq \alpha(|x(t_0)|), \qquad \forall t \geq t_0 \geq 0, \forall |x(t_0)| < c. \tag{A.8}$$

- *UAS if and only if there is a class $\mathcal{KL}$ function β and a positive constant c, independent of t_0 such that*

$$|x(t)| \leq \beta(|x(t_0)|, t - t_0), \qquad \forall t \geq t_0 \geq 0, \forall |x(t_0)| < c. \tag{A.9}$$

- *globally UAS (GUAS) if and only if inequality (A.9) is satisfied for any initial state $x(t_0)$.*

- *exponentially stable if there exists a positive constant k such that the class $\mathcal{KL}$ function β in (A.9) takes the form $\beta(r, s) = kr\exp(-\lambda s)$.*

Appendix B: Proofs

B.1 Proof of Theorem 1

We will do all proofs for the model by Wendling et al. described in Section 2.3.1, which is the most general model that encompasses the models by Stam et al. and Jansen et al. described in Sections 2.3.3 and 2.3.2 respectively. The proof for all other models can be done by first identifying the cascade structure of the observation error system, then showing that the subsystems satisfy certain properties to conclude the global exponential stability of the whole observation error system. This is performed in a similar fashion in the proof for the Wendling et al. model that follows.

From (3.1), (3.2), the dynamics of the state estimation error $\tilde{x} := x - \hat{x}$ is:

$$\dot{\tilde{x}} = A\tilde{x} + G(\gamma(Hx) - \gamma(H\hat{x} + KC\tilde{x})) + LC\tilde{x}. \tag{B.10}$$

The main idea is to consider the estimation error system (B.10) as the nominal error system (B.10) with $L = 0$ and $K = 0$ perturbed by the terms $G\gamma(H\hat{x}) - G\gamma(H\hat{x} + K(C\hat{x} - y)) + LC\tilde{x}$. As such, in Lemma 3, we build a Lyapunov function W for the nominal error system. Next, using W as a candidate Lyapunov function for (B.10), we obtain a bound for observer matrices K and L such that $\tilde{x} = 0$ is global exponential stable (GES) [85, Definition 4.5].

Lemma 3. *Consider the system* (B.10). *There exists a continuously differentiable* $W : \mathbb{R}^n \to \mathbb{R}$ *such that the following holds, for all* $\tilde{x} \in \mathbb{R}^n$*:*

$$\begin{aligned} k_1|\tilde{x}|^2 &\leq W(\tilde{x}) \leq k_2|\tilde{x}|^2, \\ \left|\tfrac{\partial W(\tilde{x})}{\partial \tilde{x}}\right| &\leq k_4|\tilde{x}|, \end{aligned}$$

and along solutions to system (B.10) *with* $L = 0$ *and* $K = 0$*:*

$$\dot{W}(\tilde{x}) \quad \leq \quad -k_3|\tilde{x}|^2, \tag{B.11}$$

where k_1*,* k_2*,* k_3 *and* k_4 *are strictly positive constants.*

Proof. The nominal estimation error system (B.10) with $L = 0$ and $K = 0$ is

$$\dot{\tilde{x}} \quad = \quad A\tilde{x} + G(\gamma(Hx) - \gamma(H\hat{x})). \tag{B.12}$$

First note that system (3.1) has solutions that are defined for all time due to the global Lipschitz property of function S and the fact that input $u \in L_\infty$ [85, Theorem 3.2]. In view of the same arguments, the solutions of system (3.2) (with y and u from system (3.1)) are also well-defined and exist for all time. Consequently, we have that (B.12) have solutions that are defined for all time.

We can decompose system (B.12) into 7 subsystems Σ_{e1} to Σ_{e7}:

$$\begin{aligned}
\Sigma_{e1}: \quad \dot{\tilde{x}}_1 &= A_1\tilde{x}_1 + \phi_1(x_{41}) - \phi_1(\hat{x}_{41}) \\
\Sigma_{e2}: \quad \dot{\tilde{x}}_2 &= A_2\tilde{x}_2 + \phi_2(x_{51}) - \phi_2(\hat{x}_{51}) \\
\Sigma_{e3}: \quad \dot{\tilde{x}}_3 &= A_3\tilde{x}_3 + \phi_3(x_{61}, x_{71}) - \phi_3(\hat{x}_{61}, \hat{x}_{71}) \\
\Sigma_{e4}: \quad \dot{\tilde{x}}_4 &= A_4\tilde{x}_4 \\
\Sigma_{e5}: \quad \dot{\tilde{x}}_5 &= A_5\tilde{x}_5 \\
\Sigma_{e6}: \quad \dot{\tilde{x}}_6 &= A_6\tilde{x}_6 \\
\Sigma_{e7}: \quad \dot{\tilde{x}}_7 &= A_7\tilde{x}_7 + \phi_7(x_{51}) - \phi_7(\hat{x}_{51}),
\end{aligned} \tag{B.13}$$

where $\tilde{x}_i := (\ \tilde{x}_{i1}, \tilde{x}_{i2}\ ) = (\ x_{i1} - \hat{x}_{i1},\ x_{i2} - \hat{x}_{i2}\ ) \in \mathbb{R}^2$ for $i = \{1, \ldots, 7\}$. $\phi_1(x_{41}) = \big(0, \theta_A a C_2 S_2(x_{41})\big)$, $\phi_2(x_{51}) = \big(0, \theta_B b C_4 S_2(x_{51})\big)$, $\phi_3(x_{61}, x_{71}) = \big(0, \theta_G g C_7 S_2(x_{61} - x_{71})\big)$, $\phi_7(x_{51}) = (0, \theta_B b C_6 S_2(x_{51}))$. Matrices $A_1, \ldots A_7$ are as defined in Section 2.4

and have eigenvalues with strictly negative real parts.

We note that subsystem Σ_{e1} is in cascade with Σ_{e4}, subsystems Σ_{e2} and Σ_{e7} are in cascade with subsystem Σ_{e5}, and subsystem Σ_{e3} is in cascade with Σ_{e6} and Σ_{e7}. Using this cascade structure, we show that the overall system (B.12) is GES by constructing the desired Lyapunov function W from the Lyapunov functions of each subsystems V_i.

We consider Lyapunov functions $V_1, \ldots, V_7$ for each subsystem Σ_{e1} to Σ_{e7} of the form:

$$V_i \quad = \quad \tilde{x}_i^T P_i \tilde{x}_i \qquad \text{for } i \in \{1, \ldots, 7\},$$

where $P_i^T = P_i > 0$ satisfies the Lyapunov equation $P_i A_i + A_i^T P_i = -I$. This is always possible as the eigenvalues of A_i have strictly negative real parts, in view of [85, Theorem 4.6]. We will show that each subsystem Σ_{ei} is input-to-state stable (ISS) using V_i [135].

For subsystems Σ_{e4}, Σ_{e5} and Σ_{e6}, taking the derivative of $V_i = \tilde{x}_i^T P_i \tilde{x}_i$ along solutions of Σ_{ei} for $i \in \{4, 5, 6\}$, we obtain:

$$\dot{V}_i \quad \leq \quad -|\tilde{x}_i|^2 \leq -\frac{1}{2}|\tilde{x}_i|^2.$$

Next, we show that V_1 is an ISS-Lyapunov function [135] for subsystem Σ_{e1} w.r.t e_4. Taking the derivative of $V_1 = \tilde{x}_1^T P_1 \tilde{x}_1$ along the solutions of Σ_{e1}, we obtain:

$$\dot{V}_1 \quad = \quad \tilde{x}_1^T (P_1 A_1 + A_1^T P_1) \tilde{x}_1 + 2\tilde{x}_1^T P_1 (\phi_1(x_{41}) - \phi_1(\hat{x}_{41})).$$

Since the function S_2 is globally Lipschitz with constant ρ_2 from (2.4), we have that:

$$|\phi_1(x_{41}) - \phi_1(\hat{x}_{41})| \quad \leq \quad \theta_A a C_2 \rho_2 |x_{41} - \hat{x}_{41}| \leq \rho_{e1} |\tilde{x}_4|,$$

where $\rho_{e1} = \theta_A a C_2 \rho_2$. Therefore,

$$\dot{V}_1 \leq -|\tilde{x}_1|^2 + 2|\tilde{x}_1||P_1|\rho_{e1}|\tilde{x}_4|.$$

Recalling that $\xi\chi \leq \frac{1}{2}\xi^2 + \frac{1}{2}\chi^2$, for any ξ, $\chi \in \mathbb{R}$ and letting $\xi = 2|P_1|\rho_{e1}|\tilde{x}_4|$ and $\chi = |\tilde{x}_1|$, we obtain the following:

$$\begin{aligned} \dot{V}_1 &\leq -|\tilde{x}_1|^2 + \frac{1}{2}|\tilde{x}_1|^2 + 2|P_1|^2\rho_{e1}^2|\tilde{x}_4|^2 \\ &\leq -\frac{1}{2}|\tilde{x}_1|^2 + 2|P_1|^2\rho_{e1}^2|\tilde{x}_4|^2. \end{aligned}$$

With similar arguments for the remaining systems, along the solutions of Σ_{ei}, for $i \in \{2, 3, 7\}$:

$$\dot{V}_i \leq -\frac{1}{2}|\tilde{x}_i|^2 + \gamma_i|\mu_i|^2,$$

where $\mu_1 = \tilde{x}_4$, $\mu_2 = \tilde{x}_5$, $\mu_3 = (\tilde{x}_6, \tilde{x}_7)$, $\mu_7 = \tilde{x}_5$ and $\gamma_i = 2|P_i|^2\rho_{ei}^2$ where $\rho_{e2} = \theta_B b C_4 \rho_2$, $\rho_{e3} = \theta_G g C_7 \rho_2$ and $\rho_{e7} = \theta_B b C_6 \rho_2$.

The composite Lyapunov function for the overall system (B.13) W is constructed using the Lyapunov function for each subsystem V_i [134]. We consider the candidate Lyapunov function that is positive definite and radially unbounded:

$$W = a_1V_1 + a_2V_2 + a_3V_3 + V_4 + V_5 + V_6 + a_7V_7, \tag{B.14}$$

for a_1, a_2, a_3, $a_7 > 0$, which are determined below.

The derivative of (B.14) along the solution of the overall error system (B.13) is

$$\begin{aligned}\dot{W} \leq\ & -a_1\frac{1}{2}|\tilde{x}_1|^2 + a_1\gamma_1|\tilde{x}_4|^2 - a_2\frac{1}{2}|\tilde{x}_2|^2 \\ & +a_2\gamma_2|\tilde{x}_5|^2 - a_3\frac{1}{2}|\tilde{x}_3|^2 + a_3\gamma_3(|\tilde{x}_6|^2 + |\tilde{x}_7|^2) \\ & -\frac{1}{2}|\tilde{x}_4|^2 - \frac{1}{2}|\tilde{x}_5|^2 - \frac{1}{2}|\tilde{x}_6|^2 - a_7\frac{1}{2}|\tilde{x}_7|^2 + a_7\gamma_7|\tilde{x}_5|^2.\end{aligned}$$

By taking $0 < a_1 < \frac{1}{2\gamma_1}$, $0 < a_2 < \frac{1}{\gamma_2}\left(\frac{1}{2} - a_7\gamma_7\right)$, $0 < a_3 < \frac{1}{2\gamma_3}$ and $a_7 > 2a_3\gamma_3$, we obtain

$$\begin{aligned}\dot{W} \leq\ & -\frac{1}{8\gamma_1}|\tilde{x}_1|^2 - \frac{1}{2\gamma_2}\left(\frac{1}{2} - \gamma_7\right)|\tilde{x}_2|^2 - \frac{1}{8\gamma_3}|\tilde{x}_3|^2 \\ & -\frac{1}{2}|\tilde{x}_4|^2 - \frac{1}{2}|\tilde{x}_5|^2 - \frac{1}{2}|\tilde{x}_6|^2 - \frac{1}{4}|\tilde{x}_7|^2. \end{aligned} \tag{B.15}$$

By letting $\tilde{P} = \text{diag}(a_1P_1,\ a_2P_2,\ a_3P_3,\ P_4,\ P_5,\ P_6,\ a_7P_7)$ as well as denoting $\lambda_{min}(\tilde{P})$ and $\lambda_{max}(\tilde{P})$ as the maximum and minimum eigenvalues of $\tilde{P}$ respectively, we have shown that (B.11) is fulfilled with $k_1 = \lambda_{min}(\tilde{P}) > 0$, $k_2 = \lambda_{max}(\tilde{P}) > 0$, $k_3 = \min\{\frac{1}{8\gamma_1}, \frac{1}{2\gamma_2}\left(\frac{1}{2} - \gamma_7\right), \frac{1}{8\gamma_3}, \frac{1}{4}\} > 0$ and $k_4 = 2|\tilde{P}| > 0$. □

The derivative of W along the solutions to (B.10) is:

$$\begin{aligned}\dot{W} &= \frac{\partial W}{\partial \tilde{x}}\Big(A\tilde{x} + G\big(\gamma(Hx) - \gamma(H\hat{x} + KC\tilde{x})\big) + LC\tilde{x}\Big), \\ &= \frac{\partial W}{\partial \tilde{x}}\Big(\underbrace{A\tilde{x} + G\gamma(Hx) - G\gamma(H\hat{x})}_{\text{nominal estimation error system (B.12)}} \\ &\qquad + \underbrace{G\gamma(H\hat{x}) - G\gamma(H\hat{x} + KC\tilde{x}) + LC\tilde{x}}_{\text{perturbation terms}}\Big)\end{aligned}$$

From Lemma 3:

$$\dot{W} \leq -k_3|\tilde{x}|^2 + \frac{\partial W}{\partial \tilde{x}}G(\gamma(H\hat{x}) - \gamma(H\hat{x} + KC\tilde{x})) + \frac{\partial W}{\partial \tilde{x}}LC\tilde{x}.$$

As γ is globally Lipschitz, $|\gamma(H\hat{x}) - \gamma(H\hat{x} + KC\tilde{x})| \leq \rho|KC\tilde{x}| \leq \rho|K||C||\tilde{x}|$. From

(B.11):

$$\begin{aligned}\dot{W} &\leq -k_3|\tilde{x}|^2 + \left|\frac{\partial W}{\partial \tilde{x}}\right| |G|\rho|K||C||\tilde{x}| + \left|\frac{\partial W}{\partial \tilde{x}}\right| |L||C||\tilde{x}| \\ &\leq -k_3|\tilde{x}|^2 + k_4|C||\tilde{x}|^2(\rho|G||K| + |L|),\end{aligned}$$

where k_3 and k_4 are constructed in Lemma 3.

Therefore, if K and L satisfy the following condition:

$$\rho|K||G| + |L| \quad < \quad \frac{k_3}{k_4|C|},$$

then

$$\dot{W} \quad \leq \quad -\tilde{k}_3|\tilde{x}|^2,$$

where $\tilde{k}_3 = k_3 - k_4|C|(\rho|G||K| + |L|) > 0$.

Therefore, the origin of the estimation error system (B.10) is GES according to [85, Definition 4.5], i.e. for all $t \geq 0$,

$$|\tilde{x}(t)| \quad \leq \quad k\exp(-\lambda t)|\tilde{x}(0)| \qquad \forall \tilde{x}(0) \in \mathbb{R}^n,$$

for $k, \lambda > 0$. □

B.2 Proof of Theorem 2

The proof that follows is performed for the Wendling at. al. model, where the models by Stam et al. as well as Jansen and Rit can be derived from. The proof for these models can be performed in a similar fashion.

From (3.5) and (3.6), the perturbed error system is:

$$\begin{aligned}\dot{\tilde{x}} &= A\tilde{x} + G(p^\star + \epsilon_p)\gamma(Hx) - G(p^\star)\gamma\Big(H\hat{x} + K\big(C\hat{x} - (y + \epsilon_y)\big)\Big) \\ &\quad -L(C\hat{x} - (y + \epsilon_y)) + \sigma(u,\, y,\, p^\star + \epsilon_p) - \sigma(u + \epsilon_u,\, y + \epsilon_y,\, p^\star) + \epsilon_{sys} \\ &= \underbrace{(A + LC)\tilde{x} + \Psi(x,\, \hat{x})}_{\text{nominal system (B.10) from Theorem 1}} + \underbrace{\Psi_\epsilon(x,\, \hat{x},\, \epsilon_y,\, \epsilon_u,\, \epsilon_p,\, \epsilon_{sys})}_{\text{perturbation terms}},\end{aligned} \tag{B.16}$$

where $K = (\kappa_1, \ldots, \kappa_m)$,
$\Psi = \Big(0,\, \theta_A a C_2\big(S_2(x_{41}) - S_2(\hat{x}_{41} - \kappa_1 C\tilde{x})\big),\, 0,\, \theta_B b C_4\big(S_2(x_{51}) - S_2(\hat{x}_{51} - \kappa_2 C\tilde{x}\big),\, 0,$
$\theta_G g C_7\big(S_2(x_{61} - x_{71}) - S_2(\hat{x}_{61} - \hat{x}_{71} - \kappa_3 C\tilde{x})\big),\, 0,\, 0,\, 0,\, 0,\, 0,\, 0,\, 0,\, \theta_B b C_6\big(S_2(x_{51}) - S_2(\hat{x}_{51} - \kappa_2 C\tilde{x})\big)\Big)$ and
$\Psi_\epsilon = \Big(0,\, \theta_A a C_2\big(S_2(\hat{x}_{41} - \kappa_1 C\tilde{x}) - S_2(\hat{x}_{41} - \kappa_1 C\tilde{x} - \epsilon_y)\big) + a C_2 S_2(x_{41})\epsilon_p + au\epsilon_p +$
$\theta_A a \epsilon_u,\, 0,\, \theta_B b C_4\big(S_2(\hat{x}_{51} - \kappa_2 C\tilde{x}) - S_2(\hat{x}_{51} - \kappa_2 C\tilde{x} - \epsilon_y)\big) + b C_4 S_2(x_{51})\epsilon_p,\, 0,$
$\theta_G g C_7\big(S_2(\hat{x}_{61} - \hat{x}_{71} - \kappa_3 C\tilde{x}) - S_2(\hat{x}_{61} - \hat{x}_{71} - \kappa_3 C\tilde{x} - \epsilon_y)\big) + g C_7 S_2(x_{61} - x_{71})\epsilon_p,\, 0,$
$\theta_A a C_1\big(S_2(y) - S_2(y + \epsilon_y)\big) + a C_1 S_2(y)\epsilon_p,\, \theta_A a C_3\big(S_2(y) - S_2(y + \epsilon_y)\big) + a C_3 S_2(y)\epsilon_\theta,\, 0,$
$\theta_A a C_5\big(S_2(y) - S_2(y + \epsilon_y)\big) + a C_5 S_2(y)\epsilon_p,\, 0,\, \theta_B b C_6\big(S_2(\hat{x}_{51} - \kappa_2 C\tilde{x}) - S_2(\hat{x}_{51} - \kappa_2 C\tilde{x} -$
$\epsilon_y)\big) + b C_6 S_2(x_{51})\epsilon_p\Big) + \epsilon_{sys}$.

We will show that the solutions of the error system (B.16) is input-to-state stable (ISS) [133] with respect to the uncertainties ϵ_y, ϵ_u, ϵ_p and ϵ_{sys}. For this purpose, we use function W as defined in (B.14).

The derivative of W along the solutions of (B.16) is:

$$\dot{W} \;=\; \frac{\partial W}{\partial \tilde{x}}\Big((A + LC)\tilde{x} + \Psi(x,\, \hat{x})\Big) + \frac{\partial W}{\partial \tilde{x}}\Psi_\epsilon(x,\, \hat{x},\, \epsilon_y,\, \epsilon_u,\, \epsilon_p,\, \epsilon_{sys}).$$

From Theorem 1, there exists $\tilde{k}_3, \tilde{k}_4 > 0$ such that:

$$\dot{W} \;\leq\; -\tilde{k}_3|\tilde{x}|^2 + \tilde{k}_4|\tilde{x}||\Psi_\epsilon|.$$

Using the fact that the function S_2 is globally Lipschitz with Lipschitz constant ρ_2 and $S_2(z) \leq \alpha_2$ for any $z \in \mathbb{R}$ as defined in (2.4), we obtain:

$$\dot{W} \;\leq\; -\tilde{k}_3|\tilde{x}|^2 + \tilde{k}_4|\tilde{x}|\big(\sigma_y(|\epsilon_y|) + \sigma_p(|\epsilon_p|) + \sigma_u(|\epsilon_u|) + |\epsilon_{sys}|\big)$$

where $\sigma_y(|\epsilon_y|) \;=\; \big(|(\theta_A a C_2\rho,\, \theta_B b C_4\rho,\, \theta_G g C_7\rho,\, \theta_A a C_1\rho,\, \theta_A a C_3\rho,\, \theta_A a C_5\rho,\, \theta_B b C_6\rho)| + |L|\big)|\epsilon_y|$, $\sigma_u(|\epsilon_u|) = \theta_A a|\epsilon_u|$ and $\sigma_p(|\epsilon_p|) = |\big((a C_2\alpha + a\|u\|_{[0,t]})|\epsilon_p|,\, b C_4\alpha|\epsilon_p|,\, g C_7\alpha|\epsilon_p|,$ $a C_1\alpha|\epsilon_p|,\, a C_3\alpha|\epsilon_p|,\, a C_5\alpha|\epsilon_p|,\, b C_6\alpha|\epsilon_p|\big)|$.

Therefore, if

$$|\tilde{x}| > \frac{2\tilde{k}_4}{\tilde{k}_3}\big(\sigma_y(|\epsilon_y|) + \sigma_u(|\epsilon_u|) + \sigma_p(|\epsilon_p|) + |\epsilon_{sys}|\big),$$

then

$$\dot{W} \leq -\frac{1}{2}\tilde{k}_3|\tilde{x}|^2. \tag{B.17}$$

From (B.17), [85, (4.49) of Theorem 4.19] is fulfilled and [85, (4.48) of Theorem 4.19] is satisfied with $\alpha_1 = k_1$ and $\alpha_2 = k_2$, where k_1 and k_2 are from Lemma 3 in B.1. Therefore, we can conclude that the error system (B.16) is ISS with respect to ϵ_y, ϵ_u, ϵ_p and ϵ_{sys} with gains $\gamma_y(s) = \frac{k_2}{k_1}\frac{2\tilde{k}_4}{\tilde{k}_3}\sigma_y(s)$, $\gamma_\theta(s) = \frac{k_2}{k_1}\frac{2\tilde{k}_4}{\tilde{k}_3}\sigma_p(s)$, $\gamma_u(s) = \frac{k_2}{k_1}\frac{2\tilde{k}_4}{\tilde{k}_3}\sigma_u(s)$ and $\gamma_{sys}(s) = \frac{k_2}{k_1}\frac{2\tilde{k}_4}{\tilde{k}_3}.s$, for $s \geq 0$ and $\gamma_y(0) = \gamma_\theta(0) = \gamma_u(0) = \gamma_{sys}(0) = 0$. Here, we have shown it for the model by Wendling et al. [155], but similar arguments apply to all other models considered in Chapter 2. □

B.3 Proof of Theorem 3

Let $V(\tilde{x}) = \tilde{x}^T P\tilde{x}$, its derivative along the solutions to (4.4) is:

$$\begin{aligned}\dot{V} =& \tilde{x}^T\big(P(A+LC) + (A+LC)^T P\big)\tilde{x} + 2\tilde{x}^T PG(p^\star)\delta(t)\eta \\ =& \begin{bmatrix}\tilde{x}\\ \delta(t)\eta\end{bmatrix}^T \begin{bmatrix}(A+LC)^T P + P(A+LC) & PG(p^\star)\\ G(p^\star)^T P & 0\end{bmatrix}\begin{bmatrix}\tilde{x}\\ \delta(t)\eta\end{bmatrix}.\end{aligned}$$

According to (4.5),

$$\begin{aligned}\dot{V} \leq& \begin{bmatrix}\tilde{x}\\ \delta(t)\eta\end{bmatrix}^T \begin{bmatrix}-\nu I & -(H+KC)^T\Lambda\\ -\Lambda(H+KC) & 2\Lambda\mathrm{diag}(\frac{1}{b_1},\dots,\frac{1}{b_m})\end{bmatrix}\begin{bmatrix}\tilde{x}\\ \delta(t)\eta\end{bmatrix}\\ =& -\nu\tilde{x}^T\tilde{x} - 2\eta^T\delta(t)\Lambda(H+KC)\tilde{x} + 2\eta^T\delta(t)\Lambda\mathrm{diag}\left(\frac{1}{b_1},\dots,\frac{1}{b_m}\right)\delta(t)\eta.\end{aligned}$$

Recall that $\eta = v - w = (H+KC)\tilde{x}$ and since Λ and δ are diagonal, $\Lambda = \Lambda^T$, $\delta(t) = \delta(t)^T$ and $\Lambda\delta(t) = \delta(t)\Lambda$,

$$\begin{aligned}\dot{V} \leq& -\nu\tilde{x}^T\tilde{x} - 2\eta^T\delta(t)\Lambda\eta + 2\eta^T\delta(t)\Lambda\mathrm{diag}\left(\frac{1}{b_1},\dots,\frac{1}{b_m}\right)\delta(t)\eta\\ =& -\nu\tilde{x}^T\tilde{x} - 2\eta^T\left(\delta(t)\Lambda - \delta(t)\Lambda\mathrm{diag}\left(\frac{1}{b_1},\dots,\frac{1}{b_m}\right)\delta(t)\right)\eta.\end{aligned}$$

If we examine the last term component-wise, we obtain:

$$\delta_i(t)\lambda_i - \delta_i(t)\lambda_i b_i^{-1}\delta_i(t) = \delta_i(t)\lambda_i(1 - b_i^{-1}\delta_i(t)).$$

From Assumption 9 and (4.3), we know that $1 - b_i^{-1}\delta_i(t) > 0$, hence,

$$\left(\delta(t)\Lambda - \delta(t)\Lambda \text{diag}\left(\frac{1}{b_1}, \ldots, \frac{1}{b_m}\right)\delta(t)\right) \geq 0.$$

We obtain $\dot{V} \leq -\nu\tilde{x}^T\tilde{x}$. By [85, Theorem 4.10], we deduce that (4.6) holds. □

B.4 Proof of Theorem 4

From (4.7) and (4.8), the observation error system is as follows:

$$\begin{aligned}\dot{\tilde{x}} &= (A+LC)\tilde{x} + L\epsilon_y + \sigma(u,y) - \sigma(u+\epsilon_u, y+\epsilon_y) \\ &\quad + G\gamma(Hx) - G\gamma\Big(H\hat{x} + K\big(C\hat{x} - (y+\epsilon_y)\big)\Big) + \epsilon_{sys}.\end{aligned} \tag{B.18}$$

By introducing $\pm G\gamma\Big(H\hat{x} + K\big(C\hat{x} - y\big)\Big)$ to (B.18), we obtain an observation error system that can be considered as the nominal system (4.4) perturbed by a term that depends on ϵ_y, ϵ_u and ϵ_{sys} as shown below:

$$\begin{aligned}\dot{\tilde{x}} &= \underbrace{(A+LC)\tilde{x} + G\big(\gamma(Hx) - \gamma(H\hat{x} + K\big(C\hat{x} - y\big)\big)\big)}_{\text{nominal system (4.4)}} \\ &\quad + \underbrace{\Psi_\epsilon(x, \hat{x}, \epsilon_y, \epsilon_u, \epsilon_{sys})}_{\text{perturbation terms}},\end{aligned} \tag{B.19}$$

where $\Psi_\epsilon = G\gamma\big(H\hat{x} + K(C\hat{x} - y)\big) - G\gamma\Big(H\hat{x} + K\big(C\hat{x} - (y+\epsilon_y)\big)\Big) + L\epsilon_y + \sigma(y,u) - \sigma(u+\epsilon_u, y+\epsilon_y) + \epsilon_{sys}$.

Let $V(e) = \tilde{x}^T P\tilde{x}$, by following the same arguments as in the proof of Theorem 3,

along solutions to (B.19), we obtain

$$\dot{V} \leq -\nu\tilde{x}^T\tilde{x} + 2\tilde{x}^T P\Psi_\epsilon. \tag{B.20}$$

The perturbation Ψ_ϵ can be bounded as follows:

$$|\Psi_\epsilon| \quad \leq \quad \sigma_y|\epsilon_y| + \sigma_u|\epsilon_u| + |\epsilon_{sys}|, \tag{B.21}$$

where $\sigma_y = |G||\text{diag}(b_1,\ldots,b_m)||K|+|L|+|\text{diag}(\tilde{b}_1,\ldots,\tilde{b}_n)|$ and $\sigma_u = |\text{diag}(\tilde{b}_1,\ldots,\tilde{b}_n)|$.

From (B.20) and the bound on the perturbation terms (B.21):

$$\dot{V} \quad \leq \quad -\nu|\tilde{x}|^2 + 2|\tilde{x}||P|\big(\sigma_y|\epsilon_y| + \sigma_u|\epsilon_u| + |\epsilon_{sys}|\big).$$

Consequently, if $|\tilde{x}| > \frac{4|P|}{\nu}\big(\sigma_y|\epsilon_y| + \sigma_u|\epsilon_u| + |\epsilon_{sys}|\big)$, then $\dot{V} \leq -\frac{\nu}{2}|\tilde{x}|^2$. By [85, Theorem 4.19], (4.9) is satisfied with $\gamma_y = \frac{\lambda_{max}(P)}{\lambda_{min}(P)}\frac{4|P|}{\nu}\sigma_y$, $\gamma_u = \frac{\lambda_{max}(P)}{\lambda_{min}(P)}\frac{4|P|}{\nu}\sigma_u$ and $\gamma_{\epsilon_{sys}} = \frac{\lambda_{max}(P)}{\lambda_{min}(P)}\frac{4|P|}{\nu}$. □

B.5 Proof of Proposition 1

Let $p^\star \in \Theta$, the perturbed model (4.10) has solutions for all $t \geq 0$ as γ and $\tilde{\sigma}$ are globally Lipschitz, the input u and the uncertainty ϵ_p are in $\mathcal{L}_\infty$ [85, Theorem 3.2]. We can write the perturbed model (4.10) in the form of (4.7) with $\epsilon_{sys} = G(p^\star+\epsilon_p)\gamma(Hx) - G(p^\star)\gamma(Hx) + \tilde{\sigma}(u,y,p^\star+\epsilon_p) - \tilde{\sigma}(y,u,p^\star)$. As Assumptions 9-10 are satisfied, we apply Theorem 4 and obtain the following:

$$|\tilde{x}(t)| \leq \bar{k}\exp(-\bar{\beta}t)|\tilde{x}(0)| + \bar{\gamma}_y\|\epsilon_y\|_{[0,t]} + \bar{\gamma}_u\|\epsilon_u\|_{[0,t]} + \gamma_{\epsilon_{sys}}\|\epsilon_{sys}\|_{[0,t]}. \tag{B.22}$$

Noting that the uncertainty is $\epsilon_{sys} = G(p^\star+\epsilon_p)\gamma(Hx) - G(p^\star)\gamma(Hx) + \tilde{\sigma}(u,y,p^\star+\epsilon_p) - \tilde{\sigma}(u,y,p^\star)$, we use the fact that γ and $\tilde{\sigma}$ are globally Lipschitz and bounded to show that ϵ_p satisfies the following

$$|\epsilon_p| \leq \sigma_p|\epsilon_p|, \tag{B.23}$$

where $\sigma_p = |(aC_2\alpha, bC_4\alpha)| + |(a\|u\|_{[0,\infty]}, aC_3\alpha, aC_1\alpha)|$, recalling that $\|u\|_{[0,\infty]} \leq 320\text{mV}$

[74, Section 3.1]. We deduce from (B.22) and (B.23) that (4.19) is satisfied with $\bar{\gamma}_p = \sigma_p \gamma_{\epsilon_{sys}}$. □

B.6 Proof of Theorem 5

Firstly, $x(t)$ exists for all $t \geq 0$ by Theorem 3.2 of [86], because γ is globally Lipschitz and u is a continuous function that is defined for all $t \geq 0$. We now show that the observation error system satisfies property (4.24) by taking the derivative of the Lyapunov function $V(\tilde{x}) = \tilde{x}^T P \tilde{x}$ along the solutions of (4.22), where $\chi = (\tilde{x}, \delta(t)\eta, w, d)$:

$$\begin{aligned}
\dot{V}(\tilde{x}) &= \tilde{x}^T \big(P(A+LC) + (A+LC)^T P\big)\tilde{x} + 2\tilde{x}^T PG\delta(t)\eta \\
&\quad -2\tilde{x}^T PLDw + 2\tilde{x}^T PBd \\
&= \chi^T \begin{bmatrix} P(A+LC) + (A+LC)^T P & PG & -PLD & PB \\ \star & 0 & 0 & 0 \\ \star & \star & 0 & 0 \\ \star & \star & \star & 0 \end{bmatrix} \chi.
\end{aligned}$$

Applying (4.25), we obtain:

$$\begin{aligned}
\dot{V}(\tilde{x}) &\leq \chi^T \begin{bmatrix} -\mathbb{I} & -(H+KC)^T M & 0 & 0 \\ \star & -\mathcal{E}(M) & MKD & 0 \\ \star & \star & \mu_w \mathbb{I} & 0 \\ \star & \star & \star & \mu_d \mathbb{I} \end{bmatrix} \chi \\
&= -|\tilde{x}|^2 - 2\tilde{x}^T (H+KC)^T M\delta(t)\eta + 2\eta^T \delta(t) MKDw \\
&\quad -\eta^T \delta(t)^T \mathcal{E}(M)\delta(t)\eta + \mu_w |w|^2 + \mu_d |d|^2.
\end{aligned}$$

Recall that $\eta := z - v = (H+KC)\tilde{x} - KDw$, hence $(H+KC)\tilde{x} = \eta + KDw$. Therefore,

$$\begin{aligned}
\dot{V}(\tilde{x}) \quad \leq \quad & -|\tilde{x}|^2 - 2(\eta + KDw)^T M\delta(t)\eta + 2\eta^T \delta(t) MKDw \\
& -\eta^T \delta(t)^T \mathcal{E}(M)\delta(t)\eta + \mu_w |w|^2 + \mu_d |d|^2.
\end{aligned}$$

Noting that $\delta(t) = \text{diag}\,(\delta_1(t), \ldots, \delta_n(t)) = \delta(t)^T$,

$$\begin{aligned}&\dot{V}(\tilde{x}) + |\tilde{x}|^2 - \mu_w|w|^2 - \mu_d|d|^2 \\ &\quad\leq -2\eta^T \left(M\delta(t) - \delta(t)M\text{diag}\left(\frac{1}{b_1}, \ldots, \frac{1}{b_m}\right)\delta(t)\right)\eta.\end{aligned}$$

We examine $M\delta(t) - \delta(t)M\text{diag}\left(\frac{1}{b_1}, \ldots, \frac{1}{b_m}\right)\delta(t)$ component by component, i.e. $\delta_i(t)m_i - \delta_i(t)^2 m_i {b_i}^{-1} = \delta_i(t)m_i\left(1 - \delta_i(t){b_i}^{-1}\right)$. As $\delta_i(t)$, $m_i > 0$ and by Assumption 9, $1 - \delta_i(t){b_i}^{-1} \geq 0$, we obtain $\delta(t)M - \delta(t)M\text{diag}\left(\frac{1}{b_1}, \ldots, \frac{1}{b_m}\right)\delta(t) \geq 0$. Hence, $\dot{V}(\tilde{x}) + |\tilde{x}|^2 - \mu_w|w|^2 - \mu_d|d|^2 \leq 0$. As explained in Chapter 4.2, this implies that the observation error system satisfies properties (4.24) as required. □

B.7 Proof of Theorem 6

Before starting the proof of Theorem 6, we first state a technical lemma that is used.

Lemma 4. *Consider system (5.5). There exist $M \in \mathbb{R}_{>0}$ such that for any $d \geq 1$, any $\mathcal{L}_\infty$ signal $\hat{x}, y, u$ the solution of $\dot{\Upsilon} = A\Upsilon + \Delta\phi(y, u, \hat{x})$ with $\Upsilon(0) = 0$, satisfies $|\Upsilon(t)| \leq M$ for all $t \geq 0$.* □

Proof. Note that $|\Delta\phi(y, u, \hat{x})| \leq |\Delta||\phi(y, u, \hat{x})| \leq |\phi(y, u, \hat{x})|$ since $d \geq 1$. Moreover, it can be verified that ϕ is upper bounded by a constant (independent of d) in view of Chapter 5.1. As a consequence, we can directly conclude that $|\Upsilon(t)|$ can be bounded by a constant M independent of d for all time since the matrix A is Hurwitz and of the initial condition $\Upsilon(0)$ as it was chosen to be $\Upsilon(0) = 0$ in the algorithm (5.5). □

We first consider the system in $\tilde{x}$ coordinate. According to (5.4) and (5.5), we have

$$\dot{\tilde{x}} = A\tilde{x} + \big(\phi(y, u, x) - \phi(y, u, \hat{x})\big)\theta + \phi(y, u, \hat{x})\tilde{p} - \Gamma(y - \hat{y}).$$

We scale the error $\tilde{x}$ as such $\bar{x} := \Delta\tilde{x}$. Noting that $\Delta A\Delta^{-1} = A$:

$$\dot{\bar{x}} = A\bar{x} + \Delta\big(\phi(y,u,x) - \phi(y,u,\hat{x})\big)p^\star + \Delta\phi(y,u,\hat{x})\tilde{p} - \Delta\Gamma(y-\hat{y}). \tag{B.24}$$

We now proceed to a dynamical change of coordinates by introducing η as follows which was as proposed in [160]:

$$\eta \;\; := \;\; \bar{x} - \Upsilon\tilde{p}.$$

Noting that $C\Delta^{-1} = dC$, the dynamics of variables η and $\tilde{p}$ are given by:

$$\left\{\begin{array}{rcl} \dot{\eta} &=& A\eta + \Delta\big(\phi(y,u,x) - \phi(y,u,\hat{x})\big)p^\star \\ \dot{\tilde{p}} &=& -dP\Upsilon^{\mathrm{T}}C^{\mathrm{T}}C(\eta + \Upsilon\tilde{p}). \end{array}\right. \tag{B.25}$$

We note that, contrary to [160], we have not obtained a nice cascade system since the η-system also depends on x and $\hat{x}$. In the following, we will invoke small-gain type arguments to conclude the stability of system (B.25). We consider the Lyapunov-type functions $V_1(\eta) = \eta^{\mathrm{T}}S\eta$ and $V_2(\tilde{p},P) = \tilde{p}^{\mathrm{T}}P^{-1}\tilde{p}$ where S is a real symmetric positive definitive matrix such that $A^{\mathrm{T}}S + SA = -\mathbb{I}$ (such a matrix always exists according to Theorem 4.6 in [86] since A is Hurwitz). It can be noted that V_2 satisfies $\lambda_1|\tilde{p}|^2 \leq V_2(\tilde{p},P) \leq \lambda_2|\tilde{p}|^2$ where λ_1,λ_2 are independent of d for $d \geq 1$ according to [162, Lemma 1], which will be useful for our purpose. We first consider V_1. Along solutions to (B.25), we have:

$$\dot{V}_1(t) = -|\eta|^2 + 2\eta^{\mathrm{T}}S\Delta\big(\phi(y,u,x) - \phi(y,u,\hat{x})\big)p^\star. \tag{B.26}$$

Noting that $\phi(x,y,u) = \big(\phi_0(y), \phi_1(x_0,u)\big)$, so $\phi(x,y,u) - \hat{\phi}(\hat{x},y,u) = (0, \phi_1(x_0,u) - \hat{\phi}_1(\hat{x}_0,u))$. As a consequence,

$$\Delta(\phi(y,u,x) - \hat{\phi}(y,u,\hat{x})) = \big(0, \tfrac{1}{d}(\phi_1(x_0,u) - \hat{\phi}_1(\hat{x}_0,u))\big)$$

from which we deduce in (B.26):

$$\begin{array}{rcl} \dot{V}_1(t) &\leq& -|\eta|^2 + 2|\eta||S||p^\star||\Delta\big(\phi(x_0,u) - \phi(\hat{x}_0,u)\big)| \\ &\leq& -|\eta|^2 + 2|\eta||S||p^\star|\big(\tfrac{1}{d}|\phi_1(x_0,u) - \phi_1(\hat{x}_0,u)|\big). \end{array}$$

By assumption, ϕ_1 is globally Lipschitz with a constant that we denote $L_1 > 0$, thus $|\phi_1(x_0, u) - \phi_1(\hat{x}_0, u)| \leq L_1|\tilde{x}_0|$. Note also that the scaling induced by the matrix Δ does not affect the $\tilde{x}_0$ part of the estimation error $\tilde{x}$, so that we have $\tilde{x}_0 = \bar{x}_0$. As a consequence,

$$\begin{aligned}\dot{V}_1(t) &\leq -|\eta|^2 + \tfrac{2}{d}|\eta||S||p^\star|L_1|\tilde{x}_0| \\ &\leq -|\eta|^2 + \tfrac{2}{d}|\eta||S||p^\star|L_1(|\eta| + |\Upsilon\tilde{p}|) \\ &= (-1 + \tfrac{2}{d}|S||p^\star|L_1)|\eta|^2 + \tfrac{2}{d}|\eta||S||p^\star|L_1|\Upsilon\tilde{p}|.\end{aligned} \tag{B.27}$$

We define

$$d^\star := \max\{1, 4L_1|p^\star||S|\}. \tag{B.28}$$

Choose d such that $d \geq d^\star$, so that $-1 + \frac{2}{d}|S||p^\star|L_1 < -\frac{1}{2}$, in that way: $\dot{V}_1(t) \leq -\frac{1}{2}|\eta|^2 + \frac{2}{d}|S|L_1|p^\star||\eta||\Upsilon\tilde{p}|$. According to Lemma 4, we always have $|\Upsilon| \leq M$, therefore: $\dot{V}_1(t) \leq -\frac{1}{2}|\eta|^2 + \frac{2}{d}|S||\theta|L_1M|\eta||\tilde{p}|$, from which we deduce, by invoking the fact that S and P^{-1} are symmetric positive definite with lower and upper bounds independent of d (see [162, Lemma 1]), that there exists $\sigma_1, \sigma_2 \in \mathbb{R}_{>0}$ independent of d such that:

$$\begin{aligned}\dot{V}_1(t) &\leq -\sigma_1 V_1 + \tfrac{1}{d}\sigma_2\sqrt{V_1}\sqrt{V_2} \\ &= -\tfrac{\sigma_1}{2}V_1 - \tfrac{\sigma_1}{2}V_1 + \tfrac{1}{d}\sigma_2\sqrt{V_1}\sqrt{V_2},\end{aligned} \tag{B.29}$$

as a consequence, noting that $\frac{1}{d}\sigma_2\sqrt{V_1}\sqrt{V_2} \leq \frac{\sigma_1}{2}V_1$ can be written as $\frac{1}{d^2}\sigma_2^2 V_2 \leq \frac{\sigma_1^2}{4}V_1$:

$$\left(\tfrac{\sigma}{d^2}V_2 \leq V_1\right) \;\Rightarrow\; \left(\dot{V}_1(t) \leq -\tfrac{\sigma_1}{2}V_1\right), \tag{B.30}$$

with $\sigma = 4\left(\frac{\sigma_2}{\sigma_1}\right)^2$. That is all we need so far concerning V_1. Consider V_2, along solutions to (5.5) and (B.25), the following is satisfied (recall that $V_2(\tilde{p}, P) = \tilde{p}^{\mathrm{T}}P^{-1}\tilde{p}$):

$$\begin{aligned}\dot{V}_2(t) &= -\tilde{p}^{\mathrm{T}}P^{-1}\dot{P}P^{-1}\tilde{p} + 2\tilde{p}^{\mathrm{T}}P^{-1}\dot{\tilde{p}} \\ &= -dV_2 - d\tilde{p}^{\mathrm{T}}\Upsilon^{\mathrm{T}}C^{\mathrm{T}}C\Upsilon\tilde{p} - 2d\tilde{p}^{\mathrm{T}}\Upsilon^{\mathrm{T}}C^{\mathrm{T}}C\eta,\end{aligned}$$

since $\tilde{p}^{\mathrm{T}}\Upsilon^{\mathrm{T}}C^{\mathrm{T}}C\Upsilon\tilde{p} \geq 0$, $\dot{V}_2(t) \leq -dV_2 - 2d\tilde{p}^{\mathrm{T}}\Upsilon^{\mathrm{T}}C^{\mathrm{T}}C\eta$. By invoking the same arguments as after (B.29) and since Υ is bounded by M according to Lemma 4, there exists $\gamma \in \mathbb{R}_{>0}$ independent of d such that $|2d\tilde{p}^{\mathrm{T}}\Upsilon^{\mathrm{T}}C^{\mathrm{T}}C\eta| \leq d\gamma\sqrt{V_1}\sqrt{V_2}$ so that:

$$\begin{aligned}\dot{V}_2(t) &\leq -dV_2 + \gamma d\sqrt{V_1}\sqrt{V_2} \\ &= -\tfrac{d}{2}V_2 - \tfrac{d}{2}V_2 + \gamma d\sqrt{V_1}\sqrt{V_2}\end{aligned}$$

since $\gamma d\sqrt{V_1}\sqrt{V_2} \leq \frac{d}{2}V_2$ is equivalent to $\gamma^2 V_1 \leq \frac{1}{4}V_2$, we deduce that:

$$\Big(\bar{\sigma}V_1 \leq V_2\Big) \quad \Rightarrow \quad \Big(\dot{V}_2(t) \leq -\tfrac{d}{2}V_2\Big), \tag{B.31}$$

with $\bar{\sigma} = 4\gamma^2$. Following the proof of Theorem 3.1 in [76], we are going to define a Lyapunov function for the system (B.25) based on V_1 and V_2. Let d be sufficiently large such that $\frac{\sigma}{d^2} < \bar{\sigma}^{-1}$ (i.e. such that the small gain condition is satisifed) and take $\chi > 0$ such that:

$$\tfrac{\sigma}{d^2} < \chi < \bar{\sigma}^{-1}. \tag{B.32}$$

We introduce the candidate Lyapunov function, $V = \max\{V_1, \chi V_2\}$. Function V is the maximum of two continuously differentiable functions, therefore it is locally Lipschitz and so differentiable almost everywhere according to Rademacher's theorem. Along solutions to (B.25), when $V = V_1$ that means $V_1 \geq \chi V_2 \geq \frac{\sigma}{d^2}V_2$. (B.30) shows us that $\dot{V}(t) \leq -\frac{\sigma_1}{2}V(t)$ almost everywhere. Similarly, when $V = \chi V_2$ that means $V_2 \geq \chi^{-1}V_1 \geq \bar{\sigma}V_1$ so (B.31) shows us that $\dot{V}(t) \leq -\frac{d}{2}\chi V(t)$ almost everywhere. As a consequence, we have that, almost everywhere: $\dot{V}(t) \leq -\min\{\frac{\sigma_1}{2}, \frac{d}{2}\chi\}V$. Taking $d \geq d^\star \geq 1$,

$$\dot{V}(t) \leq -\min\{\tfrac{\sigma_1}{2}, \tfrac{d}{2}\chi\}V \leq -\min\{\tfrac{\sigma_1}{2}, \tfrac{1}{2}\chi\}V =: -\lambda V.$$

Using the comparison principle (see Lemma 3.4 in [86]) and the fact that V is positive definite and radially unbounded, we obtain:

$$|(\eta(t), \tilde{p}(t))| \quad \leq \quad \bar{\beta}\Big(|(\eta(0), \tilde{p}(0))|, t\Big) \qquad \forall t \geq 0, \tag{B.33}$$

where $\bar{\beta} \in \mathcal{KL}$. By the definition of $\tilde{x}$,

$$\begin{aligned} |(\tilde{x}, \tilde{p})| &= |(\Delta^{-1}\eta + \Delta^{-1}\Upsilon\tilde{p}, \tilde{p})| \\ &\leq d|(\eta + \Upsilon\tilde{p}, \tilde{p})| \leq d|\eta + \Upsilon\tilde{p}| + d|\tilde{p}| \\ &\leq d|\eta| + d(1+M)|\tilde{p}| \leq d(1+M)|(\eta, \tilde{p})|. \end{aligned}$$

From (B.33) and $\eta(0) = \Delta\tilde{x}(0) - \Upsilon(0)\tilde{p}(0)$. As $\Upsilon(0) = 0$ by initialisation, we have that $\eta(0) = \Delta\tilde{x}(0)$. Recalling that $|\Delta| \leq 1$ (since $d \geq 1$), hence $|\eta(0)| \leq |\Delta||\tilde{x}(0)| \leq |\tilde{x}(0)|$

and we obtain:

$$\begin{aligned}|(\tilde{x}(t),\tilde{p}(t))| &\leq d(1+M)\bar{\beta}\Big(|(\eta(0),\tilde{p}(0))|,t\Big)\\ &\leq d(1+M)\bar{\beta}\Big(|(\tilde{x}(0),\tilde{p}(0))|,t\Big).\end{aligned}$$

Therefore, we obtain the desired property (5.8) with $\beta_d(s,t) = d(1+M)\bar{\beta}(s,t)$. □

B.8 Proof of Proposition 2

We write the state error system (6.4), for all $i \in \{1,\dots,N\}$ in the following form

$$\begin{aligned}\dot{\xi} &= F_i(\xi,\varrho,p^\star,u,x)\\ \eta &= H(\xi,x,\varrho,p^\star).\end{aligned} \tag{B.34}$$

where $\xi \in \mathbb{R}^{n_x}$, $\varrho \in \mathbb{R}^{n_p}$ and $\eta \in \mathbb{R}^{n_y}$. We denote the solution of system (B.34) at time t with parameter error ϱ as $\xi(t,\varrho)$, where we have omitted its dependance on its initial condition $x(0)$ and on its input u. Similarly, we denote the output of (B.34) as $\eta(t,\varrho)$.

Let $W(t,\varrho) := \int_{t-T_f}^{t} |\eta(s,\varrho)|^2 ds$. We show that W is continuous in ϱ. First, since the function f is continuously differentiable, $\varrho \in \tilde{\Theta}$, $p^\star \in \Theta$, $u \in \mathcal{M}_{\Delta_u}$ and $x \in B_{K_x}$ (by Assumption 16), F_i is locally Lipschitz in ξ, uniformly in ϱ, $p^\star$, u and x by Lemma 2.3 of [85]. Hence, $\xi(t,\varrho)$ is continuous in t and ϱ by Theorem 2.6 in [85]. As a consequence, we deduce that $\eta(t,\varrho)$ is also continuous in ϱ and t by using the fact that H is continuous in view of Assumption 15. Thus, we have that $\rho \mapsto |\eta(t,\rho)|$ is continuous in $\tilde{\Theta}$. Moreover, $|\eta(t,\varrho)| \leq \phi(t)$, with $\phi(t) := \max_{\varrho\in\tilde{\Theta}} |\eta(t,\varrho)|^2$ integrable on $[t-T,t]$ using the fact that $\tilde{\Theta}$ is a compact set. Hence, we deduce that W is continuous in ϱ on $\tilde{\Theta}$, by using the dominated convergence theorem [14, Theorem 10.27].

So far, for any fixed t, $W(t,\varrho) \geq 0$ for all ϱ by definition and we have established that $W(t,\varrho)$ is continuous in ϱ. Hence, we can lower bound the function W by a positive definite and continuous function $\tilde{W}$ in ϱ. This allows us to apply Lemma 3.5 in [85] on $\tilde{W}$ to obtain a class $\mathcal{K}$ function $\alpha_{\tilde{y}}$ such that $\tilde{W}(t,\varrho) \geq \alpha_{\tilde{y}}(|\varrho|)$. Since, $W(t,\varrho) \geq \tilde{W}(t,\varrho)$, (6.9) holds as desired. □

B.9 Proof of Lemma 1

We first prove in Claims 1 and 2 desirable properties of the state error systems (6.4) and the monitoring signals μ_i, where $i \in \{1, \dots, N\}$. In Claim 1, the state error system (6.4) is shown to be locally ISS (see [85, Definition 5.2]) with respect to the parameter error $\tilde{p}_i$. This property, along with Assumption 18, are the key ingredients that allow us to show in Claim 2 that the monitoring signals are lower and upper bounded by functions of the parameter error $\tilde{p}_i$ that are strictly increasing with $|\tilde{p}_i|$. These properties allow us to conclude the proof of Lemma 1. Claims 1 and 2 are also later used in the proof of Theorem 7.

Claim 1. *Consider the state error system (6.3) for $i \in \{1, \dots, N\}$ under Assumptions 15-17. There exist constants $\bar{k}$, $\bar{\lambda} > 0$ such that for any $\Delta_{\tilde{x}}$, Δ_x, $\Delta_u > 0$, there exists a class $\mathcal{K}$ function $\bar{\gamma}_{\tilde{x}}$ such that for all $(\tilde{x}_i(0), x(0)) \in B_{\Delta_{\tilde{x}}} \times B_{\Delta_x}$, for all $\tilde{p}_i \in \tilde{\Theta}$ and for all $u \in \mathcal{M}_{\Delta_u}$, the corresponding solution satisfies*

$$|\tilde{x}_i(t)| \leq \bar{k} \exp(-\bar{\lambda} t)|\tilde{x}_i(0)| + \bar{\gamma}_{\tilde{x}}(|\tilde{p}_i|), \qquad \forall t \geq 0. \tag{B.35}$$

□

Proof of Claim 1. For $i \in \{1, \dots, N\}$, by Assumption 17, the following holds for all $\tilde{x}_i \in \mathbb{R}^{n_x}$, $x \in \mathbb{R}^{n_x}$, $u \in \mathbb{R}^{n_u}$, $p^\star \in \Theta$ and $\tilde{p}_i \in \tilde{\Theta}$

$$a_1|\tilde{x}_i|^2 \leq V_i(\tilde{x}_i) \leq a_2|\tilde{x}_i|^2, \tag{B.36}$$

$$\frac{\partial V_i}{\partial \tilde{x}_i} F_i(\tilde{x}_i, \tilde{p}_i, p^\star, u, x) \leq -\lambda_0 V_i(\tilde{x}_i) + \tilde{\gamma}(\tilde{p}_i, x, u). \tag{B.37}$$

Given Δ_x, $\Delta_u > 0$, let $K_x > 0$ be generated by Assumption 16 such that $|x(t)| \leq K_x$, for all $t \geq 0$ and $u \in \mathcal{M}_{\Delta_u}$. Since $\tilde{\gamma}$ is a continuous function and $\tilde{\gamma}(0, z, \bar{z}) = 0$, for all $z \in \mathbb{R}^{n_x}$, $\bar{z} \in \mathbb{R}^{n_u}$, we have that

$\tilde{\gamma}_1(\tilde{p}_i) := \max_{|x| \leq K_x, |u| \leq \Delta_u} \tilde{\gamma}(\tilde{p}_i, x, u)$ is a continuous and positive definite function in

$\tilde{p}_i$. By Lemma 3.5 in [85], there exists a class $\mathcal{K}$ function $\bar{\gamma}$ such that for $\tilde{p}_i \in \tilde{\Theta}$

$$\tilde{\gamma}_1(\tilde{p}_i) \leq \bar{\gamma}(|\tilde{p}_i|), \tag{B.38}$$

as $\tilde{\Theta}$ is a compact set.

We use (B.36), (B.37) and (B.38) to obtain the following for all $|x| \leq K_x$ and $|u| \leq \Delta_u$

$$\frac{\partial V_i}{\partial \tilde{x}_i} F_i(\tilde{x}_i, \tilde{p}_i, p^\star, u, x) \leq -\lambda_0 V_i(\tilde{x}_i) + \bar{\gamma}(|\tilde{p}_i|). \tag{B.39}$$

By the comparison principle (Lemma 2.5 in [85]), for all $(x(0), \tilde{x}_i(0)) \in B_{\Delta_x} \times B_{\Delta_{\tilde{x}}}$, $u \in \mathcal{M}_{\Delta_u}$, the corresponding solutions of (B.39) are

$$V_i(\tilde{x}_i(t)) \leq \exp(-\lambda_0 t) V_i(\tilde{x}_i(0)) + \frac{1}{\lambda_0} \bar{\gamma}(|\tilde{p}_i|), \tag{B.40}$$

for $t \in [0, t_{\max})$, where $t_{\max} \in \mathbb{R}_{\geq 0} \cup \infty$ is the solution's maximal time of existence. We use (B.36) and the fact that for any a, $b > 0$, $\sqrt{a+b} \leq \sqrt{2a} + \sqrt{2b}$ to obtain (B.35) for all $t \in [0, t_{\max})$ as desired with $\bar{k} := \sqrt{\frac{2a_2}{a_1}}$, $\bar{\lambda} := \frac{\lambda_0}{2}$ and $\bar{\gamma}_{\tilde{x}}(r) := \sqrt{\frac{2}{a_1 \lambda_0}} \sqrt{\bar{\gamma}(r)}$. We deduce from (B.40) that $t_{\max} = \infty$ by contradiction. □

Claim 2. *Consider the plant (6.1), the state error system (6.4) and the monitoring signal (6.7), for $i \in \{1, \ldots, N\}$ under Assumptions 15-18. For any $\Delta_{\tilde{x}}$, Δ_x, $\Delta_u > 0$, there exist class $\mathcal{K}$ functions $\underline{\chi}$ and $\bar{\chi}$ and for any $\epsilon > 0$, there exists a constant $T = T(\Delta_{\tilde{x}}, \Delta_x, \Delta_u, \epsilon) > 0$, such that the monitoring signal μ_i in (6.7) satisfy the following for all $(x(0), \tilde{x}_i(0)) \in B_{\Delta_x} \times B_{\Delta_{\tilde{x}}}$, for some $u \in \mathcal{M}_{\Delta_u}$ that satisfies Assumption 18, for any $p^\star \in \Theta$ and for all $\tilde{p}_i \in \tilde{\Theta}$*

$$\underline{\chi}(|\tilde{p}_i|) + c_\mu \leq \mu_i(t) \leq \bar{\chi}(|\tilde{p}_i|) + \epsilon + c_\mu, \qquad \forall t \geq T. \tag{B.41}$$

□

Proof of Claim 2. Given $\Delta_{\tilde{x}}$, Δ_x, Δ_u, $\epsilon > 0$, we first construct the following from Assumptions 15-18.

- Let constants $\bar{k}, \bar{\lambda} > 0$ and a class $\mathcal{K}$ function $\bar{\gamma}_{\tilde{x}}$ be generated from Claim 1. Given that for $i \in \{1, \ldots, N\}$, $\tilde{p}_i \in \tilde{\Theta}$, where $\tilde{\Theta}$ is a compact set, there exists $K_{\tilde{p}} > 0$ such that $|\tilde{p}_i| \leq K_{\tilde{p}}$. By Claim 1, the solutions of (6.4) satisfy the following for all $(\tilde{x}_i(0), x(0)) \in B_{\Delta_{\tilde{x}}} \times B_{\Delta_x}$, for all $(p^\star, \tilde{p}_i) \in \Theta \times \tilde{\Theta}$, for all $u \in \mathcal{M}_{\Delta_u}$ and for all $t \geq 0$

$$|\tilde{x}_i(t)| \leq \bar{k} \exp(-\bar{\lambda} t)|\tilde{x}_i(0)| + \bar{\gamma}_{\tilde{x}}(|\tilde{p}_i|) \leq \bar{k}\Delta_{\tilde{x}} + \bar{\gamma}_{\tilde{x}}(K_{\tilde{p}}) =: K_{\tilde{x}}. \tag{B.42}$$

- Since h is continuously differentiable in view of Assumption 15, h is locally Lipschitz. Therefore, there exist constants $l_{\tilde{x}}, l_{\tilde{p}} > 0$ such that for $x \in B_{K_x}$, $\tilde{x}_i \in B_{K_{\tilde{x}}}$, and $\tilde{p}_i \in \tilde{\Theta}$, $p^\star \in \Theta$, the following holds

$$\begin{aligned} |H(\tilde{x}_i, x, \tilde{p}_i, p^\star) - H(0, x, 0, p^\star)| &= |H(\tilde{x}_i, x, \tilde{p}_i, p^\star)| \\ &= |h(\tilde{x}_i + x, \tilde{p}_i + p^\star) - h(x, p^\star)| \\ &\leq l_{\tilde{x}}|\tilde{x}_i| + l_{\tilde{p}}|\tilde{p}_i|, \end{aligned} \tag{B.43}$$

  where $H(0, x, 0, p^\star) = 0$ in view of the definition of H in (6.4).

- Let $\lambda > 0$ come from (6.7) and given $\epsilon > 0$, we choose $\epsilon_{\tilde{x}}$, $\epsilon_\mu > 0$ sufficiently small such that

$$\epsilon_\mu + \frac{4 l_{\tilde{x}}^2 \epsilon_{\tilde{x}}^2}{\lambda} = \epsilon. \tag{B.44}$$

  Let $T_{\tilde{x}} > 0$ be such that for all $|x(0)| \leq \Delta_x$, $|\tilde{x}_i(0)| \leq \Delta_{\tilde{x}}$ and $u \in \mathcal{M}_{\Delta_u}$

$$\bar{k} \exp(-\bar{\lambda} t)|\tilde{x}_i(0)| \leq \epsilon_{\tilde{x}}, \qquad \forall t \geq T_{\tilde{x}}. \tag{B.45}$$

Let $T_\mu \geq T_{\tilde{x}}$ be such that

$$\frac{1}{\lambda}\exp(-\lambda(T_\mu - T_{\tilde{x}}))(l_{\tilde{x}}K_{\tilde{x}} + l\tilde{p}K_{\tilde{p}})^2 \leq \epsilon_\mu. \tag{B.46}$$

- Let the constant $T_f > 0$ and the class $\mathcal{K}$ function $\alpha_{\tilde{y}}$ be generated by Assumption 18 given $\Delta_{\tilde{x}}$, Δ_x and $\Delta_u > 0$ such that for all $(\tilde{x}_i(0), x(0)) \in B_{\Delta_{\tilde{x}}} \times B_{\Delta_x}$, for some $u \in \mathcal{M}_{\Delta_u}$, for all $\tilde{p}_i \in \tilde{\Theta}$, the solution to (6.4) satisfies the following inequality

$$\int_{t-T_f}^{t} |H(\tilde{x}_i(s), x(s), \tilde{p}_i, p^\star)|^2 ds \geq \alpha_{\tilde{y}}(|\tilde{p}_i|), \qquad \forall t \geq T_f. \tag{B.47}$$

- With the functions and constants above, we define the following

$$T \geq \max\{T_\mu, T_{\tilde{x}}, T_f\}, \tag{B.48}$$

$$\underline{\chi}(r) := \exp(-\lambda T_f)\alpha_{\tilde{y}}(r), \tag{B.49}$$

$$\bar{\chi}(r) := \frac{4l_{\tilde{x}}^2}{\lambda}\bar{\gamma}_{\tilde{x}}^2(r) + \frac{2l_{\tilde{p}}^2}{\lambda}r^2. \tag{B.50}$$

Note that the class $\mathcal{K}$ functions $\underline{\chi}$ and $\bar{\chi}$ depend only on $\Delta_{\tilde{x}}$, Δ_x, Δ_u and not on ϵ. Let $t \geq T$ where T is defined in (B.48). By definition of the monitoring signals in (6.7), for $i \in \{1, \ldots, N\}$

$$\mu_i(t) = \int_0^t \exp(-\lambda(t-s))|H(\tilde{x}_i(s), x(s), \tilde{p}_i, p^\star)|^2 ds + c_\mu. \tag{B.51}$$

We first establish the lower bound for μ_i by considering the intervals $[0, t - T_f]$ and

$[t - T_f, t]$ to obtain

$$\begin{aligned}\mu_i(t) &= \int_0^{t-T_f} \exp(-\lambda(t-s))|H(\tilde{x}_i(s), x(s), \tilde{p}_i, p^\star)|^2 ds \\ &\quad + \int_{t-T_f}^{t} \exp(-\lambda(t-s))|H(\tilde{x}_i(s), x(s), \tilde{p}_i, p^\star)|^2 ds + c_\mu \\ &\geq \int_{t-T_f}^{t} \exp(-\lambda(t-s))|H(\tilde{x}_i(s), x(s), \tilde{p}_i, p^\star)|^2 ds + c_\mu. \end{aligned} \tag{B.52}$$

As $s \mapsto \exp(\lambda s)$ is strictly increasing, we obtain the following lower bound

$$\mu_i(t) \geq \exp(-\lambda T_f) \int_{t-T_f}^{t} |H(\tilde{x}_i(s), x(s), \tilde{p}_i, p^\star)|^2 ds + c_\mu. \tag{B.53}$$

From (B.47), (B.49) and our choice of T, since $t \geq T \geq T_f$

$$\mu_i(t) \geq \exp(-\lambda T_f)\alpha_{\tilde{y}}(|\tilde{p}_i|) + c_\mu = \underline{\chi}(|\tilde{p}_i|) + c_\mu. \tag{B.54}$$

We now obtain the upper bound of μ_i by considering the integration intervals $[0, T_{\tilde{x}}]$ and $[T_{\tilde{x}}, t]$, since $t \geq T \geq T_{\tilde{x}}$.

$$\begin{aligned}\mu_i(t) &= \int_0^{T_{\tilde{x}}} \exp(-\lambda(t-s))|H(\tilde{x}_i(s), x(s), \tilde{p}_i, p^\star)|^2 ds \\ &\quad + \int_{T_{\tilde{x}}}^{t} \exp(-\lambda(t-s))|H(\tilde{x}_i(s), x(s), \tilde{p}_i, p^\star)|^2 ds + c_\mu. \end{aligned} \tag{B.55}$$

Using (B.43) and the fact that for any a, $b \geq 0$, $(a+b)^2 \leq 2a^2 + 2b^2$, we obtain

$$\begin{aligned}
\mu_i(t) &\leq \frac{1}{\lambda}\exp(-\lambda(t-T_{\tilde{x}}))\left(\sup_{s\in[0,T_{\tilde{x}}]}|H(\tilde{x}_i(s),x(s),\tilde{p}_i,p^\star)|^2\right) \\
&\quad + \int_{T_{\tilde{x}}}^t \exp(-\lambda(t-s))\Big(l_{\tilde{x}}|\tilde{x}_i(s)| + l_{\tilde{p}}|\tilde{p}_i|\Big)^2 ds + c_\mu \\
&\leq \frac{1}{\lambda}\exp(-\lambda(t-T_{\tilde{x}}))\left(\sup_{s\in[0,T_{\tilde{x}}]}|H(\tilde{x}_i(s),x(s),\tilde{p}_i,p^\star)|^2\right) \\
&\quad + \int_{T_{\tilde{x}}}^t \exp(-\lambda(t-s))\Big(2l_{\tilde{x}}^2|\tilde{x}_i(s)|^2 + 2l_{\tilde{p}}^2|\tilde{p}_i|^2\Big) ds + c_\mu. \qquad \text{(B.56)}
\end{aligned}$$

We also have from Claim 1 and (B.45) that for all $i \in \{1,\ldots,N\}$

$$|\tilde{x}_i(t)| \leq \epsilon_{\tilde{x}} + \gamma_{\tilde{x}}(|\tilde{p}_i|), \qquad \forall t \geq T_{\tilde{x}}, \qquad \text{(B.57)}$$

which implies that, using the fact that for any a, $b > 0$, $(a+b)^2 \leq 2a^2 + 2b^2$

$$|\tilde{x}_i(t)|^2 \leq 2\epsilon_{\tilde{x}}^2 + 2\gamma_{\tilde{x}}^2(|\tilde{p}_i|), \qquad \forall t \geq T_{\tilde{x}}. \qquad \text{(B.58)}$$

Also, by (B.42) and (B.43), we obtain

$$\begin{aligned}
\sup_{s\in[0,T_{\tilde{x}}]}|H(\tilde{x}_i(s),x(s),\tilde{p}_i,p^\star)|^2 &\leq \sup_{s\in[0,T_{\tilde{x}}]}(l_{\tilde{x}}|\tilde{x}_i(s)| + l_{\tilde{p}}|\tilde{p}_i|)^2 \\
&= (l_{\tilde{x}}K_{\tilde{x}} + l_{\tilde{p}}K_{\tilde{p}})^2. \qquad \text{(B.59)}
\end{aligned}$$

Hence, from (B.56) and in view of (B.46), (B.58) and (B.59), we have the following for $t \geq T \geq T_\mu \geq T_{\tilde{x}}$

$$\begin{aligned}
\mu_i(t) &\leq \epsilon_\mu + c_\mu + \int_{T_{\tilde{x}}}^t \exp(-\lambda(t-s))\Big(4l_{\tilde{x}}^2\epsilon_{\tilde{x}}^2 + 4l_{\tilde{x}}^2\gamma_{\tilde{x}}^2(|\tilde{p}_i|) + 2l_{\tilde{p}}^2|\tilde{p}_i|^2\Big) ds \\
&= \epsilon_\mu + c_\mu + \left(\int_{T_{\tilde{x}}}^t \exp(-\lambda(t-s)) ds\right)\Big(4l_{\tilde{x}}^2\epsilon_{\tilde{x}}^2 + 4l_{\tilde{x}}^2\gamma_{\tilde{x}}^2(|\tilde{p}_i|) + 2l_{\tilde{p}}^2|\tilde{p}_i|^2\Big). \qquad \text{(B.60)}
\end{aligned}$$

Note that $\int_{T_{\tilde{x}}}^{t} \exp(-\lambda(t-s))ds = \frac{1}{\lambda}\big(1-\exp(-\lambda(t-T_{\tilde{x}}))\big)$ and for $t \geq T_{\tilde{x}}$, we have that $\int_{T_{\tilde{x}}}^{t} \exp(-\lambda(t-s))ds \leq \frac{1}{\lambda}$. Therefore, from (B.60), for $t \geq T_{\tilde{x}}$

$$\mu_i(t) \quad \leq \quad \epsilon_\mu + c_\mu + \frac{4l_{\tilde{x}}^2\epsilon_{\tilde{x}}^2}{\lambda} + \frac{4l_{\tilde{x}}^2}{\lambda}\gamma_{\tilde{x}}^2(|\tilde{p}_i|) + \frac{2l_{\tilde{p}}^2}{\lambda}|\tilde{p}_i|^2. \tag{B.61}$$

By the definition of ϵ and $\bar{\chi}$ in (B.44) and (B.50) respectively:

$$\mu_i(t) \leq \epsilon + \bar{\chi}(|\tilde{p}_i|) + c_\mu. \tag{B.62}$$

Therefore, from (B.54) and (B.62), we complete the proof.

□

We are now ready to prove Lemma 1. Given that the switching logic (6.12) is scale-independent, the switching signal σ is not affected if the monitoring signal $\mu_i(t)$, $i \in \{1, \ldots, N\}$ is replaced by its scaled version, for $t \geq 0$

$$\bar{\mu}_i(t) := \exp(\lambda t)\mu_i(t) = \exp(\lambda t)c_\mu + \int_0^t \exp(\lambda s)|\tilde{y}_i(s)|^2 ds. \tag{B.63}$$

Note that the scaled monitoring signal $\bar{\mu}_i$ has the following properties

1. $\bar{\mu}_i(0) = c_\mu > 0$.
2. $\bar{\mu}_i(t)$ is non-decreasing.

The scaled monitoring signal $\bar{\mu}_i$ is then used for analysis purposes in [67] and [95, Chapter 6]. It has been shown in [151] that in the supervisory control of nonlinear plants with disturbances, the scale independent hysteresis switching logic σ in (6.12) satisfy dwell time conditions for bounded disturbances. We follow this approach in the analysis that follows. By Lemma 4.2 in [151], we have that the number of switches in the interval $[0, t)$, $N_\sigma(t, 0)$ satisfies the following

$$N_\sigma(t,0) \leq N + \frac{N}{\ln(1+h)} \ln\left(\frac{\bar{\mu}_l(t)}{\min_{q\in\{1,\ldots,N\}} \bar{\mu}_q(0)}\right) \tag{B.64}$$

For an arbitrary ϵ, Δ_x, $\Delta_{\tilde{x}}$, $\Delta_u > 0$, let a class $\mathcal{K}$ function $\bar{\chi}$ and a constant

$T = T(\Delta_x, \Delta_{\tilde{x}}, \Delta_u, \epsilon) > 0$ be generated by Claim 2. From Claim 2, for all $\varrho \in \tilde{\Theta}$, $(x(0), \xi(0)) \in B_{\Delta_x} \times B_{\Delta_{\tilde{x}}}$, for all $u \in \mathcal{M}_{\Delta_u}$ that satisfies Assumption 18 and for all $t \geq T$

$$\mu_i(t) \leq \bar{\chi}(|\varrho|) + \epsilon + c_\mu. \tag{B.65}$$

Hence, by definition (B.63)

$$\bar{\mu}_i(t) \leq \exp(\lambda t)\Big(\bar{\chi}(|\varrho|) + \epsilon + c_\mu\Big). \tag{B.66}$$

From (B.64), (B.66) and $\bar{\mu}_i(0) = \exp(\lambda 0)c_\mu = c_\mu$, we obtain the following for all $t \geq T$

$$\begin{aligned} N_\sigma(t,0) &\leq N + \frac{N}{\ln(1+h)} \ln\left(\frac{\exp(\lambda t)\Big(\bar{\chi}(|\varrho|) + \epsilon + c_\mu\Big)}{\exp(\lambda 0)c_\mu}\right) \\ &= N + \frac{N}{\ln(1+h)} \ln\left(\frac{\bar{\chi}(|\varrho|) + \epsilon}{c_\mu} + 1\right) + \frac{t}{\left(\frac{\ln(1+h)}{\lambda N}\right)}. \end{aligned} \tag{B.67}$$

Recall from (6.13) that a switching signal σ satisfies average dwell time conditions. It holds that

$$N_\sigma(t,0) \leq N_0 + \frac{t}{\tau_a}. \tag{B.68}$$

Therefore, comparing (B.67) and (B.68) yields

$$\tau_a = \tfrac{\ln(1+h)}{\lambda N}, \quad N_0 = N + \tfrac{N}{\ln(1+h)} \ln\left(\tfrac{\bar{\chi}(|\varrho|)+\epsilon}{c_\mu} + 1\right) \geq 1. \tag{B.69}$$

Hence, from (B.69), we have shown that σ has an average dwell time by obtaining (6.14) as desired. □

B.10 Proof of Theorem 7

The proof of Theorem 7 uses Claims 1-2 from the proof of Lemma 1 (Appendix B.9). They show properties of the state error systems (6.4) and the monitoring signals (6.7) which enable us to complete the proof. The state estimation error system (6.17) switches between the individual state estimation error systems (6.4) in a manner that

does not affect the individual systems (6.4) as explained in Remark 8. Therefore,

$$|\tilde{x}_{\sigma(t)}(t)| \leq \max_{i\in\{1,\ldots,N\}} |\tilde{x}_i(t)|, \qquad \forall t \geq 0. \tag{B.70}$$

By Claim 1, for $i \in \{1,\ldots,N\}$, for all $\tilde{x}_i(0) \in B_{\Delta_{\tilde{x}}}$, $\tilde{p}_i \in \tilde{\Theta}$ and $u \in \mathcal{M}_{\Delta_u}$, the corresponding solution of (6.4) satisfies

$$|\tilde{x}_i(t)| \leq \bar{k}\exp(-\bar{\lambda}t)|\tilde{x}_i(0)| + \bar{\gamma}_{\tilde{x}}(|\tilde{p}_i|). \tag{B.71}$$

As $\tilde{p}_i \in \tilde{\Theta}$ for $i \in \{1,\ldots,N\}$, where $\tilde{\Theta}$ is a compact set, let the constant $K_{\tilde{p}} > 0$ be such that $|\tilde{p}_i| \leq K_{\tilde{p}}$. We have from (B.71) and (B.77) that

$$|\tilde{x}_i(t)| \leq \bar{k}\Delta_{\tilde{x}} + \bar{\gamma}_{\tilde{x}}(K_{\tilde{p}}) =: \bar{K}_{\tilde{x}}, \qquad \forall t \geq 0. \tag{B.72}$$

Hence, we have that the solution of the switched state error system (6.17) satisfies

$$|\tilde{x}_{\sigma(t)}(t)| \leq \bar{K}_{\tilde{x}}, \qquad \forall t \geq 0. \tag{B.73}$$

Hence, $\tilde{x}_{\sigma(t)}(t)$ cannot escape to infinity in finite time. Moreover, we have proved in Lemma 1 that the switching logic σ satisfies average dwell time conditions. Therefore, no Zeno behaviour can occur. Hence, $\tilde{x}_{\sigma(t)}(t)$ is defined for all $t \geq 0$.

We now construct the core elements for the rest of the proof.

- Given $\Delta_{\tilde{x}}$, Δ_x, $\Delta_u > 0$, we generate constants $\bar{k}$, $\bar{\lambda} > 0$ and a class $\mathcal{K}$ function $\bar{\gamma}_{\tilde{x}}$ from Claim 1. Let $\nu_1 > 0$ be such that

$$\nu_1 \leq \gamma_{\tilde{x}}^{-1}(\nu_{\tilde{x}}), \tag{B.74}$$

  where $\nu_{\tilde{x}} > 0$ is given.

- Let class $\mathcal{K}$ functions $\bar{\chi}$ and $\underline{\chi}$ be generated by Claim 2. Choose ϵ, $h^\star > 0$ sufficiently small be such that

$$\frac{1}{1+h^\star}\underline{\chi}(\min\{\nu_{\tilde{p}}, \nu_1\}) - \frac{h^\star}{1+h^\star}c_\mu - \epsilon > 0, \tag{B.75}$$

where $\nu_{\tilde{p}} > 0$ and $c_\mu > 0$ (from the monitoring signal (6.7)) are given and let $T > 0$ be generated by Claim 2 given the chosen ϵ from (B.75). Let

$$\delta^\star := \bar{\chi}^{-1}\left(\frac{1}{1+h^\star}\underline{\chi}(\min\{\nu_{\tilde{p}}, \nu_1\}) - \frac{h^\star}{1+h^\star}c_\mu - \epsilon\right). \tag{B.76}$$

Note that $\delta^\star > 0$ due to (B.75).

- As $\tilde{p}_i \in \tilde{\Theta}$ for $i \in \{1, \ldots, N\}$, where $\tilde{\Theta}$ is a compact set, let the constant $K_{\tilde{p}} > 0$ be such that $|\tilde{p}_i| \leq K_{\tilde{p}}$. We define

$$\bar{K}_{\tilde{x}} := \bar{k}\Delta_{\tilde{x}} + \bar{\gamma}_{\tilde{x}}(K_{\tilde{p}}). \tag{B.77}$$

Let $t \geq 0$. Recall that we denote the chosen monitoring signal at each time t as $\mu_{\sigma(t)}(t)$. There are two possible scenarios: (i) $\mu_{\sigma(t)}(t) = \min_{j\in\{1,\ldots,N\}} \mu_j(t)$ and (ii) $\mu_{\sigma(t)}(t) \neq \min_{j\in\{1,\ldots,N\}} \mu_j(t)$. We first consider case (ii), since case (i) immediately follows as we will see. The switching condition of σ defined in (6.12) is not satisfied, which means that for all $q \in \{1, \ldots, N\}$,

$$\mu_q(t)(1+h) \geq \mu_{\sigma(t)}(t). \tag{B.78}$$

Consequently, by Claim 2, for all $t \geq T$

$$(1+h)\Big(\bar{\chi}(|\tilde{p}_q|) + \epsilon + c_\mu\Big) \geq \underline{\chi}(|\tilde{p}_{\sigma(t)}(t)|) + c_\mu. \tag{B.79}$$

As $h \in (0, h^\star]$, $(1+h^\star)\Big(\bar{\chi}\left(\min_{q\in\{1,\ldots,N\}} |\tilde{p}_q|\right) + \epsilon + c_\mu\Big) - c_\mu \geq \underline{\chi}(|\tilde{p}_{\sigma(t)}(t)|)$.
Since $\min_{q\in\{1,\ldots,N\}} |\tilde{p}_q| \leq \delta^\star$,

$$(1+h^\star)\Big(\bar{\chi}\left(\delta^\star\right) + \epsilon + c_\mu\Big) - c_\mu \geq \underline{\chi}(|\tilde{p}_{\sigma(t)}(t)|). \tag{B.80}$$

By the definition of $\delta^\star$ in (B.76), we obtain the following from (B.80)

$$\begin{aligned}\underline{\chi}^{-1}\Big((1+h^\star)\big(\bar{\chi}(\delta^\star) + \epsilon + c_\mu\big) - c_\mu\Big) &= \underline{\chi}^{-1}\Big((1+h^\star)\big(\bar{\chi}(\delta^\star) + \epsilon\big) + h^\star c_\mu\Big) \\ &\geq |\tilde{p}_{\sigma(t)}(t)|.\end{aligned}$$

Hence,

$$\begin{aligned} |\tilde{p}_{\sigma(t)}(t)| &\leq \underline{\chi}^{-1}\Big((1+h^\star)\big(\tfrac{1}{1+h^\star}\underline{\chi}(\min\{\nu_{\tilde{p}},\nu_1\}) - \tfrac{h^\star}{1+h^\star}c_\mu\big) - \epsilon + \epsilon + h^\star c_\mu\Big) \\ &= \min\{\nu_1,\nu_{\tilde{p}}\}. \end{aligned} \tag{B.81}$$

We now consider (i) where $\mu_{\sigma(t)}(t) = \min_{j\in\{1,\dots,N\}}\mu_j(t)$. For all $q \in \{1,\dots,N\}$

$$\mu_{\sigma(t)}(t) \leq \mu_q(t). \tag{B.82}$$

Since $h > 0$, (B.82) implies (B.78) . Therefore, we conclude the following from (B.81)

$$|\tilde{p}_{\sigma(t)}(t)| \leq \min\{\nu_1,\nu_{\tilde{p}}\} \leq \nu_{\tilde{p}}, \qquad \forall t \geq T. \tag{B.83}$$

Returning to the chosen state estimation error $\tilde{x}_{\sigma(t)}(t)$, we have from (B.70) and (B.71) that for all $\tilde{x}_{\sigma(0)}(0) \in B_{\Delta_{\tilde{x}}}$

$$|\tilde{x}_{\sigma(t)}(t)| \leq \max_{i\in\{1,\dots,N\}} \bar{k}\exp(-\bar{\lambda}t)|\tilde{x}_i(0)| + \bar{\gamma}_{\tilde{x}}(|\tilde{p}_i|), \qquad \forall t \geq 0. \tag{B.84}$$

Furthermore, we have from (B.83) that the parameter error $\tilde{p}_{\sigma(t)}$ converges to the ball centered at 0 and of radius $\nu_{\tilde{p}}$ in finite time, i.e. for all $t \geq T$, $\sigma(t) \in \mathcal{S}$, where $\mathcal{S} := \{i \in \{1,\dots,N\} : |\tilde{p}_i| \leq \min\{\nu_{\tilde{p}},\nu_1\}\}$. Therefore, for all $t \geq T$

$$|\tilde{x}_{\sigma(t)}(t)| \leq \max_{i\in\mathcal{S}}|\tilde{x}_i(t)|. \tag{B.85}$$

Consequently, we have from (B.71) that

$$\begin{aligned} \limsup_{t\to\infty}|\tilde{x}_{\sigma(t)}(t)| &\leq \limsup_{t\to\infty}\max_{i\in\mathcal{S}}|\tilde{x}_i(t)| \leq \max_{i\in\mathcal{S}}\bar{\gamma}_{\tilde{x}}(|\tilde{p}_i|) && \text{(B.86)} \\ &\leq \bar{\gamma}_{\tilde{x}}(\min\{\nu_{\tilde{p}},\nu_1\}) \leq \bar{\gamma}_{\tilde{x}}(\nu_1). && \text{(B.87)} \end{aligned}$$

By (B.74), we obtain the following

$$\limsup_{t\to\infty}|\tilde{x}_{\sigma(t)}(t)| \leq \bar{\gamma}_{\tilde{x}} \circ \bar{\gamma}_{\tilde{x}}^{-1}(\nu_{\tilde{x}}) \leq \nu_{\tilde{x}}. \tag{B.88}$$

Finally, we have shown that (6.21) in view of (B.73), (B.83) and (B.88). □

B.11 Proof of Proposition 3

For $i \in \{1, \dots, N\}$, let $V_i = \tilde{x}_i^T P_i \tilde{x}_i$, where $P_i = P_i^T > 0$. Note that (6.5) is satisfied with $a_1 := \min_{i\in\{1,\dots,N\}} \lambda_{\min}(P_i)$ and $a_2 := \max_{i\in\{1,\dots,N\}} \lambda_{\max}(P_i)$. The derivative of V_i along the solutions of (6.24) is

$$\begin{aligned}\dot{V}_i(\tilde{x}_i) &= \tilde{x}_i^T \Big(P_i\big(A(p_i) + L_i(p_i)C(p_i)\big) + \big(A(p_i) + L_i(p_i)C(p_i)\big)^T P_i \Big)\tilde{x}_i \\ &\quad + 2\tilde{x}_i^T P_i \Big(\big(\tilde{A}(p_i, p^\star) + L_i(p_i)\tilde{C}(p_i, p^\star)\big)x + \tilde{B}(p_i, p^\star)u \Big). \end{aligned} \tag{B.89}$$

Let $\bar{w}(p_i, p^\star, x, u) = \big(\tilde{A}(p_i, p^\star) + L_i(p_i)\tilde{C}(p_i, p^\star)\big)x + \tilde{B}(p_i, p^\star)u$,

$$\begin{aligned}\dot{V}_i(\tilde{x}_i) &= \begin{bmatrix} \tilde{x}_i \\ \bar{w} \end{bmatrix}^T \\ &\quad \times \begin{bmatrix} P_i\big(A(p_i) + L_i(p_i)C(p_i)\big) + \big(A(p_i) + L_i(p_i)C(p_i)\big)^T P_i & P_i \\ P_i & 0 \end{bmatrix} \begin{bmatrix} \tilde{x}_i \\ \bar{w} \end{bmatrix}. \end{aligned}$$

Applying (6.25), we obtain

$$\dot{V}_i(\tilde{x}_i) \leq -\nu_i |\tilde{x}_i|^2 + \mu_i |\bar{w}|^2. \tag{B.90}$$

By Assumption 15, we have that there exists $l_c > 0$ such that $|\tilde{C}(p_i, p^\star)| \leq l_c |\tilde{p}_i|$. Also, by Assumption 19, $A(p)$ and $B(p)$ are continuous in p on Θ and since Θ is a compact set, A and B are uniformly continuous on Θ. Hence, A and B admit $l_a : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ and $l_b : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ as moduli of continuity respectively [41], i.e. $|A(p_i) - A(p^\star)| \leq l_a(|\tilde{p}_i|)$ and $|B(p_i) - B(p^\star)| \leq l_b(|\tilde{p}_i|)$, for all $p_i \in \Theta_o$ and $p^\star \in \Theta$. Consequently,

$$|\bar{w}| \leq \tilde{w}(\tilde{p}_i, x, u), \tag{B.91}$$

where $\tilde{w}(\tilde{p}_i, x, u) = l_a(|\tilde{p}_i|)|x| + l_c|L_i(p_i)||\tilde{p}_i||x| + l_b(|\tilde{p}_i|)|u|$. Therefore, from (B.90), inequality (6.6) is satisfied with $\lambda_0 := \min_{i\in\{1,\dots,N\}} \frac{\nu_i}{\max_{i\in\{1,\dots,N\}} \lambda_{\max}(P_i)} > 0$ and $\tilde{\gamma}(\tilde{p}_i, x, u) := \max_{i\in\{1,\dots,N\}} \mu_i \tilde{w}(\tilde{p}_i, x, u)^2$. Therefore, Assumption 17 holds. □

B.12 Proof of Proposition 5

Let $i \in \{1, \dots, N\}$ and denote the state error $\tilde{x}_i := \hat{x}_i - x$ and parameter error $\tilde{p}_i = p_i - p^\star$ as in Section 6.2, we obtain the following state estimation error system from (6.26) and (6.28)

$$\begin{aligned} \dot{\tilde{x}}_i &= \Big(A(p_i) + L_i(p_i)C(p_i)\Big)\tilde{x}_i + G(p_i)\Big(\gamma(w_i) - \gamma(v)\Big) + \tilde{G}(p_i, p^\star)\gamma(v) \\ &\quad + \Big(\tilde{A}(p_i, p^\star) + L_i(p_i)\tilde{C}(p_i, p^\star)\Big)x + \tilde{B}(p_i, p^\star)\sigma(u, y), \end{aligned} \tag{B.92}$$

where $v := Hx$, $w_i := H\hat{x}_i + K_i(p_i)(C(p_i)\hat{x}_i - y)$, $\tilde{A}(p_i, p^\star) := A(p_i) - A(p^\star)$, $\tilde{B}(p_i, p^\star) := B(p_i) - B(p^\star)$, $\tilde{G}(p_i, p^\star) := G(p_i) - G(p^\star)$ and $\tilde{C}(p_i, p^\star) := C(p_i) - C(p^\star)$.

Since γ_k, $k \in \{1, \dots, n_\gamma\}$ is globally Lipschitz with Lipschitz constant b_{γ_k}, there exists a time-varying gain $\delta(t) = \mathrm{diag}(\delta_1(t), \dots, \delta_{n_\gamma}(t))$, where $\delta_k(t)$ take values in the interval $[0, b_{\gamma_k}]$ so that, for $\gamma = (\gamma_1, \dots, \gamma_{n_\gamma})$

$$\gamma(w_i) - \gamma(v) = \delta(t)(w_i - v), \qquad \forall w_i, v \in \mathbb{R}^{n_\gamma},\ i \in \{1, \dots, N\}. \tag{B.93}$$

Therefore, (B.92) can be written as

$$\begin{aligned} \dot{\tilde{x}}_i &= \tilde{f}(\tilde{x}_i, p_i, t) + \tilde{g}_i(x, u, \tilde{p}_i) \\ \tilde{y}_i &= C\tilde{x}_i, \end{aligned} \tag{B.94}$$

where $\tilde{f}(\tilde{x}_i, p_i, t) := \big(A(p_i) + L_i(p_i)C(p_i)\big)\tilde{x}_i + G(p_i)\delta(t)\big(H + K_i(p_i)C(p_i)\big)\tilde{x}_i$, $\tilde{g}_i(x, u, \tilde{p}_i) := K_i(p_i)\tilde{C}(p_i, p^\star)x + \tilde{G}(p_i, p^\star)\gamma(Hx) + \big(\tilde{A}(p_i, p^\star) + L_i(p_i)\tilde{C}(p_i, p^\star)\big)x + \tilde{B}(p_i, p^\star)\sigma(u, y)$.

We choose a quadratic function $V_i(\tilde{x}_i) = \tilde{x}_i^T P_i \tilde{x}_i$, where $P_i = P_i^T > 0$. Note that V_i satisfies inequality (6.5) of Assumption 17 with $a_1 = \lambda_{min}(P_i)$ and $a_2 = \lambda_{max}(P_i)$. By following the proof of Theorem 2 in [34] with the vector $\chi_i := (\tilde{x}_i, \delta(t)(H + K_i(p_i)C)\tilde{x}_i, \bar{w})$, where $\bar{w} = \bar{w}(p_i, p^\star, x, u) := K_i(p_i)\tilde{C}(p_i, p^\star)x + \tilde{G}(p_i, p^\star)\gamma(Hx) + \big(\tilde{A}(p_i, p^\star) + L_i(p_i)\tilde{C}(p_i, p^\star)\big)x + \tilde{B}(p_i, p^\star)\sigma(u, y)$, we obtain

$$\dot{V}_i(\tilde{x}_i) \leq -\nu_i|\tilde{x}_i|^2 + \mu_i|\bar{w}|^2. \tag{B.95}$$

By Assumption 15, there exists $l_c > 0$ such that for all $p^\star \in \Theta$ and $p_i \in \Theta_o$, we have that $|\tilde{C}(p_i, p^\star)| \leq l_c|\tilde{p}_i|$. Also, by Assumption 20, $A(p)$, $B(p)$ and $G(p)$ are continuous in p on Θ. Since Θ is a compact set, A, B and G are uniformly continuous on Θ. Hence,

A, B and G admit $l_a : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, $l_b : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ and $l_g : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ as moduli of continuity respectively [41], i.e. $|A(p_i) - A(p^\star)| \leq l_a(|\tilde{p}_i|)$, $|B(p_i) - B(p^\star)| \leq l_b(|\tilde{p}_i|)$ and $|G(p_i) - G(p^\star)| \leq l_g(|\tilde{p}_i|)$, for all $p_i \in \Theta_o$ and $p^\star \in \Theta$. Therefore,

$$|\bar{w}| \leq \tilde{w}(\tilde{p}_i, x, u), \tag{B.96}$$

where $\tilde{w}(\tilde{p}_i, x, u) := l_c|K_i(p_i)||x||\tilde{p}_i| + l_g(|\tilde{p}_i|)|\gamma(Hx)| + \Big(l_a(|\tilde{p}_i|) + |L_i(p_i)|l_c|\tilde{p}_i|\Big)|x| + l_b(|\tilde{p}_i|)|\sigma(u, y)|$. By (B.95) and (B.96), inequality (6.6) of Assumption 17 is satisfied with $\lambda_0 = \frac{\min_{i\in\{1,\ldots,N\}} \nu_i}{\max_{i\in\{1,\ldots,N\}} \lambda_{max}(P_i)}$ and $\tilde{\gamma}(z_1, z_2, z_3) := \max_{i\in\{1,\ldots,N\}} \mu_i \tilde{w}(z_1, z_2, z_3)^2$. Therefore, (6.6) holds. □

Appendix C: Values and descriptions of standard constants used in the models

Parameter	Description	Standard value
$\frac{1}{a_1}$, $\frac{1}{a_2}$	Average time constant in the excitatory feedback loop	$a_1 = 55$, $a_2 = 605\,s^{-1}$
$\frac{1}{b_1}$, $\frac{1}{b_2}$	Average time constant in the inhibitory feedback loop	$b_1 = 27.5$, $b_2 = 55\,s^{-1}$
V_1, α_1, r_1	Parameters for the sigmoid function. α_1 is the maximum firing rate. r_1 is the slope of the sigmoid and V_1 is the threshold of the population's mean membrane potential.	$V_1 = 6\,\mathrm{mV}$, $\alpha_1 = 5\,\mathrm{s}^{-1}$, $r_1 = 0.56\,\mathrm{mV}^{-1}$
C_3, C_4	Average number of synaptic contacts in the inhibitory feedback loop	$C_3 = 32$, $C_4 = 3$
θ_A, θ_B	Synaptic gain of the excitatory and inhibitory populations respectively	$\theta_A = 1.65$, $\theta_B = 32$

Table 1: Standard constants used in the model by Stam et al. in [138].

Parameter	Description	Standard value
$\frac{1}{a}$	Average time constant in the excitatory feedback loop	$a = 100\,s^{-1}$
$\frac{1}{b}$	Average time constant in the slow inhibitory feedback loop	$b = 50\,s^{-1}$
V_2, α_2, r_2	Parameters for the sigmoid function. α_2 is the maximum firing rate. r_2 is the slope of the sigmoid and V_2 is the threshold of the population's mean membrane potential.	$V_2 = 6\,\mathrm{mV}$, $\alpha_2 = 5\,\mathrm{s}^{-1}$, $r_2 = 0.56\,\mathrm{mV}^{-1}$
C_1, C_2	Average number of synaptic contacts in the excitatory feedback loop	With $C = 135$, $C_1 = C$ and $C_2 = 0.8\,C$
C_3, C_4	Average number of synaptic contacts in the slow inhibitory feedback loop	$C_3 = C_4 = 0.25\,C$
θ_A and θ_B	Synaptic gain of the excitatory and inhibitory populations respectively	$\theta_A = 3.25$ and $\theta_B = 22$

Table 2: Standard constants used in the model by Jansen and Rit in [74].

Parameter	Description	Standard value
$\frac{1}{a}$	Average time constant in the excitatory feedback loop	$a = 100\,s^{-1}$
$\frac{1}{b}$	Average time constant in the slow inhibitory feedback loop	$b = 50\,s^{-1}$
$\frac{1}{g}$	Average time constant in the fast inhibitory feedback loop	$g = 500\,s^{-1}$
V_2, α_2, r_2	Parameters for the sigmoid function. α_2 is the maximum firing rate. r_2 is the slope of the sigmoid and V_2 is the threshold of the population's mean membrane potential.	$V_2 = 6\,\mathrm{mV}$, $\alpha_2 = 5\,\mathrm{s}^{-1}$, $r_2 = 0.56\,\mathrm{mV}^{-1}$
C_1, C_2	Average number of synaptic contacts in the excitatory feedback loop	With $C = 135$, $C_1 = C$ and $C_2 = 0.8\,C$
C_3, C_4	Average number of synaptic contacts in the slow inhibitory feedback loop	$C_3 = C_4 = 0.25\,C$
C_5, C_6	Average number of synaptic contacts between the fast and slow inhibitory feedback loop	$C_5 = 0.3\,C$ and $C_6 = 0.1\,C$
C_7	Average number of synaptic contacts in the fast inhibitory feedback loop	$C_7 = 0.8\,C$
θ_A, θ_B and θ_G	Synaptic gain of the excitatory, fast inhibitory and slow inhibitory populations respectively	See [155] for values corresponding to different brain activity

Table 3: Standard constants used in the model by Wendling et al. in [155]

References

[1] Medtronic receives health canada license for deep brain stimulation therapy in refractory epilepsy patients. http://wwwp.medtronic.com/Newsroom/NewsReleaseDetails.do?itemId=1331755279978&lang=en_US. Accessed: 13/02/2013. 1

[2] Responsive neurostimulation for the treatment of epilepsy, neuropace inc. product overview. http://www.neuropace.com/product/overview.html. Accessed: 13/02/2013. 1

[3] *Fact Sheet No. 166*. World Health Organization, 2001. 1

[4] *Neurological Disorders: Public Health Challenges*. World Health Organization, 2006. 1

[5] *Fact Sheet No. 999*. World Health Organization, 2012. 1

[6] M. Abbaszadeh and H.J. Marquez. A robust observer design method for continuous-time Lipschitz nonlinear systems. In *45th IEEE Conference on Decision and Control*, pages 3795–3800. IEEE, 2007. 9

[7] C. Aboky, G. Sallet, and J.C. Vivalda. Observers for Lipschitz non-linear systems. *International Journal of control*, 75(3):204–212, 2002. 9

[8] V. Adetola and M. Guay. Finite-time parameter estimation in adaptive control of nonlinear systems. *IEEE Transactions on Automatic Control*, 53(3):807–811, 2008. 88

[9] A.P. Aguiar. Multiple-model adaptive estimators: Open problems and future directions. In *Proceedings of the European Control Conference*, 2007. 10

[10] A.P. Aguiar, M. Athans, and A. Pascoal. Convergence properties of a continuous-time multiple-model adaptive estimator. In *Proceedings of the European Control Conference*, 2007. 10

[11] A.P. Aguiar, V. Hassani, A.M. Pascoal, and M. Athans. Identifcation and convergence analysis of a class of continuous-time multiple-model adaptive estimators. In *Proceedings of the 17th IFAC World Congress*, 2008. 10

[12] B.D.O. Anderson and J.B. Moore. *Optimal filtering*. Prentice-Hall Englewood Cliffs, NJ, 1979. 8, 10

[13] V. Andrieu, L. Praly, and A. Astolfi. High gain observers with updated gain and homogeneous correction terms. *Automatica*, 45(2):422 – 428, 2009. 9, 10, 30

[14] T.M. Apostol and I. Makai. *Mathematical analysis*, volume 2. Addison-Wesley Reading, MA, 1974. 114

[15] M. Arcak, H. Gorgun, L.M. Pedersen, and S. Varigonda. A nonlinear observer design for fuel cell hydrogen estimation. *IEEE Transactions on control systems technology*, 12(1):101, 2004. 9

[16] M. Arcak and P. Kokotović. Nonlinear observers: a circle criterion design and robustness analysis. *Automatica*, 37(12):1923–1930, 2001. 9, 10, 31, 32, 33, 34, 36, 37, 50, 51, 52, 55, 61, 84

[17] M. Arcak and P. Kokotović. Observer-based control of systems with slope-restricted nonlinearities. *IEEE Transactions on Automatic Control*, 46:1, 2001. 9, 10, 31, 50, 51, 52, 58, 59, 61, 83, 85

[18] J. Back and J.H. Seo. An algorithm for system immersion into nonlinear observer form: SISO case. *Automatica*, 42(2):321–328, 2006. 9

[19] J.P. Barbot, D. Boutat, and T. Floquet. An observation algorithm for nonlinear systems with unknown inputs. *Automatica*, 45(8):1970–1974, 2009. 9

[20] G. Bastin and M.R. Gevers. Stable adaptive observers for nonlinear time-varying systems. *IEEE Transactions on Automatic Control*, 33(7):650 –658, jul. 1988. 10

[21] G. Battistelli, JP Hespanha, and P. Tesi. Supervisory control of switched nonlinear systems. *International Journal of Adaptive Control and Signal Processing*, 2012. 10, 73

[22] G. Besançon. Remarks on nonlinear adaptive observer design. *Systems & control letters*, 41(4):271–280, 2000. 10

[23] G. Besançon. *Nonlinear observers and applications*. Lecture notes in control and information sciences. Springer, 2007. 8, 9, 31, 82

[24] G. Besançon, Q. Zhang, and H. Hammouri. High-gain observer based state and parameter estimation in nonlinear systems. In *Proceedings of the 6th IFAC Symposium*, 2004. 70

[25] G. Besançon, Q. Zhang, and H. Hammouri. High-gain observer based state and parameter estimation in nonlinear systems. *Nonlinear Control Systems 2004*, page 327, 2005. 31

[26] N. Boizot, E. Busvelle, and J.P. Gauthier. Brief paper: An adaptive high-gain observer for nonlinear systems. *Automatica (Journal of IFAC)*, 46(9):1483–1488, 2010. 9, 10, 31

[27] K. Busawon, M. Farza, and H. Hammouri. Observer design for a special class of nonlinear systems. *International Journal of Control*, 71(3):405–418, 1998. 9

[28] K.K. Busawon and M. Saif. A state observer for nonlinear systems. *IEEE Transactions on Automatic Control*, 44(11):2098 –2103, November 1999. 9

[29] E. Busvelle and JP Gauthier. High-gain and non high-gain observers for nonlinear systems. *Contemporary Trends in Nonlinear Geometric Control Theory,(World Scientific, Anzaldo-Meneses, Bonnard, Gauthier, Monroy-Perez, editors)*, pages 257–286, 2002. 9, 30

[30] N. Chakravarthy, S. Sabesan, L. Iasemidis, and K. Tsakalis. Controlling synchronization in a neuron-level population model. *International journal of neural systems*, 17(02):123–138, 2007. 91

[31] N. Chakravarthy, S. Sabesan, K. Tsakalis, and L. Iasemidis. Controlling epileptic seizures in a neural mass model. *Journal of Combinatorial Optimization*, 17(1):98–116, 2009. 91

[32] Y.M. Cho and R. Rajamani. A systematic approach to adaptive observer synthesis for nonlinear systems. *IEEE Transactions on Automatic Control*, 42(4):535, 1997. 10

[33] M.S. Chong, R. Postoyan, D. Nešić, L. Kuhlmann, and A. Varsavsky. Estimating the unmeasured membrane potential of neuronal populations from the eeg using a class of deterministic nonlinear filters. *Journal of Neural Engineering*, 9:026001, 2012. 64

[34] M.S. Chong, R. Postoyan, D. Nešić, L. Kuhlmann, and A. Varsavsky. A robust circle criterion observer with application to neural mass models. *Automatica*, 48:2986–2989, 2012. 64, 74, 83, 85, 127

[35] M.S. Chong, R. Postoyan, D. Nešić, L. Kuhlmann, and A. Varsavsky. A circle criterion observer for estimating the unmeasured membrane potential of neuronal populations. In *Proceedings of the Australian Control Conference*, 2011. 50, 59, 64

[36] M.S. Chong, R. Postoyan, D. Nešić, L. Kuhlmann, and A. Varsavsky. A nonlinear estimator for the activity of neuronal populations in the hippocampus. In *Proceedings of the 18th IFAC World Congress*, pages 9899–9904, 2011. 37

[37] O. David, S.J. Kiebel, L.M. Harrison, J. Mattout, J.M. Kilner, and K.J. Friston. Dynamic causal modeling of evoked responses in eeg and meg. *NeuroImage*, 30(4):1255–1272, 2006. 9

[38] P. Dayan, L.F. Abbott, and L. Abbott. *Theoretical neuroscience: Computational and mathematical modeling of neural systems*. Taylor & Francis, 2001. 14

[39] G. Deco, V. Jirsa, P.A. Robinson, M. Breakspear, and K. Friston. The dynamic brain: From spiking neurons to neural masses and cortical fields. *Cerebral Cortex*, 4(8):1–35, 2008. 7, 17, 28, 90

[40] F. Deza, D. Bossanne, E. Busvelle, J.P. Gauthier, and D. Rakotopara. Exponential observers for nonlinear systems. *IEEE Transactions on Automatic Control*, 38(3):482 –484, March 1993. 9

[41] A.V. Efimov. *Encyclopedia of Mathematics*. Springer, 2001. 126, 128

[42] L. El Ghaoui and S.I. Niculescu. *Advances in linear matrix inequality methods in control*, volume 2. Society for Industrial Mathematics, 2000. 60

[43] R. Engel and G. Kreisselmeier. A continuous-time observer which converges in finite time. *IEEE Transactions on Automatic Control*, 47(7):1202–1204, 2002. 9

[44] D. Fairhurst, I. Tyukin, H. Nijmeijer, and C. van Leeuwen. Observers for canonic models of neural oscillators. *Mathematical Modelling of Natural Phenomena*, 5(2):146–184, 2010. 9

[45] X. Fan and M. Arcak. Observer design for systems with multivariable monotone nonlinearities. *Systems & Control Letters*, 50(4):319 – 330, 2003. 9, 10, 31, 50, 52, 59, 86

[46] M. Farza, M. M'Saad, T. Maatoug, and M. Kamoun. Adaptive observers for nonlinearly parameterized class of nonlinear systems. *Automatica*, 45(10):2292–2299, 2009. 10, 31, 64, 65, 69, 70, 71, 74, 85, 86, 93

[47] M. Farza, M. M'saad, and L. Rossignol. Observer design for a class of MIMO nonlinear systems* 1. *Automatica*, 40(1):135–143, 2004. 9

[48] M. Farza, M. M'Saad, M. Triki, and T. Maatoug. High gain observer for a class of non-triangular systems. *Systems & Control Letters*, 2010. 9

[49] K.N. Fountas and J.R. Smith. *Operative Neuromodulation*, volume 2, chapter A novel closed-loop stimulation system in the control of focal, medically refractory epilepsy, pages 357–362. Springer, 2007. 91

[50] W.J. Freeman. Models of the dynamics of neural populations. *Electroencephalography and clinical neurophysiology. Supplement*, (34):9, 1978. 8, 18

[51] D.R. Freestone, P. Aram, M. Dewar, K. Scerri, D.B. Grayden, and V. Kadirkamanathan. A data-driven framework for neural field modeling. *NeuroImage*, pages 1043–1058, 2011. 9

[52] P. Frogerais. *Model and identification in epilepsy: from neuronal population dynamics to EEG signals (in French)*. PhD thesis, University of Rennes 1, 2008. 9

[53] J.P. Gauthier and G. Bornard. Observability for any u(t) of a class of nonlinear systems. *IEEE Transactions on Automatic Control*, 26(4):922 – 926, August 1981. 9

[54] J.P. Gauthier, H. Hammouri, and S. Othman. A simple observer for nonlinear systems applications to bioreactors. *IEEE Transactions on Automatic Control*, 37(6):875–880, 1992. 9, 10, 30

[55] J.P. Gauthier and D. Kazakos. Observability and observers for non-linear systems. In *The proceedings of the 25th IEEE conference on decision and control*, volume 25, pages 953 –956, 1986. 9

[56] J. Gotman. Automatic detection of seizures and spikes. *Journal of Clinical Neurophysiology*, 16(2):130, 1999. 2

[57] H.F. Grip, A. Saberi, and T.A. Johansen. Estimation of states and parameters for linear systems with nonlinearly parameterized perturbations. *Systems & Control Letters*, 2011. 10

[58] L. Grüne, E.D. Sontag, and F.R. Wirth. Asymptotic stability equals exponential stability, and iss equals finite energy gain—if you twist your eyes. *Systems & Control Letters*, 38(2):127–134, 1999. 77

[59] Y. Guan and M. Saif. A novel approach to the design of unknown input observers. *IEEE Transactions on Automatic Control*, 36(5):632–635, 1991. 9

[60] H. Hammouri and M. Farza. Nonlinear observers for locally uniformly observable systems. *ESAIM: Control, Optimisation and Calculus of Variations*, 9:353–370, 2003. 9

[61] H. Hammouri and Z. Tmar. Unknown input observer for state affine systems: A necessary and sufficient condition. *Automatica*, 46(2):271–278, 2010. 9, 35

[62] M. Hartman, N. Bauer, and A. Teel. Robust finite-time parameter estimation using a hybrid systems framework. *IEEE Transactions on Automatic Control*, 57:2956–2962, 2012. 88

[63] J.P. Hespanha, D. Liberzon, and A.S. Morse. Bounds on the number of switchings with scale-independent hysteresis: applications to supervisory control. In *Proceedings of the 39th IEEE Conference on Decision and Control*, volume 4, pages 3622–3627, 2000. 80

[64] J.P. Hespanha, D. Liberzon, and A.S. Morse. Hysteresis-based switching algorithms for supervisory control of uncertain systems. *Automatica*, 39(2):263–272, 2003. 10, 73, 79

[65] J.P. Hespanha, D. Liberzon, and A.S. Morse. Overcoming the limitations of adaptive control by means of logic-based switching. *Systems & Control Letters*, 49(1):49–65, 2003. 74

[66] J.P. Hespanha, D. Liberzon, A.S. Morse, B.D.O. Anderson, T.S. Brinsmead, and F. De Bruyne. Multiple model adaptive control. part 2: switching. *International journal of robust and nonlinear control*, 11(5):479–496, 2001. 73

[67] J.P. Hespanha and A.S. Morse. Scale-independent hysteresis switching. *Hybrid systems: computation and control*, pages 117–122, 1999. 121

[68] A. Hodgkin and A. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology*, 117:500–544, 1952. 17

[69] L.D. Iasemidis. Epileptic seizure prediction and control. *IEEE Transactions on Biomedical Engineering*, 50(5):549–558, 2003. 3

[70] S. Ibrir. Circle-criterion observers for dynamical systems with positive and non-positive slope nonlinearities. In *Proceedings of the American Control Conference*, pages 260–265. IEEE, 2007. 9

[71] P.A. Ioannou and J. Sun. *Robust adaptive control*. Prentice Hall, 1996. 69, 78

[72] A. Isidori. *Nonlinear control systems*. Springer Verlag, 1995. 30

[73] B.H. Jansen, A.B. Kavaipatti, and O. Markusson. Evoked potential enhancement using a neurophysiologically-based model. *Methods of information in medicine*, 40(4):338–345, 2001. 35

[74] B.H. Jansen and V.G. Rit. Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biological Cybernetics*, 73:357–366, 1995. 8, 12, 18, 22, 35, 55, 56, 65, 66, 68, 70, 72, 86, 91, 109, 130

[75] A.H. Jazwinski. *Stochastic processes and filtering theory*. Academic Press, 1970. 8

[76] Z.P. Jiang, I.M.Y. Mareels, and Y. Wang. A lyapunov formulation of the nonlinear small-gain theorem for interconnected ISS systems. *Automatica*, 32(8):1211–1215, 1996. 113

[77] V.K. Jirsa and H Haken. Field theory of electromagnetic brain activity. *Physical Review Letters*, 77:960–963, 1996. 90

[78] R. Kalman and R. Bucy. New results in linear prediction and filtering theory. *Trans. AMSE J. Basic Eng*, pages 95–108, 1961. 9

[79] E.R. Kandel, J.H. Schwartz, T.M. Jessell, et al. *Principles of neural science*, volume 4. McGraw-Hill New York, 2000. 14

[80] I. Karafyllis and Z.P. Jiang. Hybrid dead-beat observers for a class of nonlinear systems. *Systems & Control Letters*, 60(8):608–617, 2011. 88

[81] D. Karagiannis and A. Astolfi. Nonlinear observer design using invariant manifolds and applications. In *44th IEEE Conference on Decision and Control and European Control Conference. CDC-ECC'05.*, pages 7775–7780. IEEE, 2006. 9

[82] D. Karagiannis, M. Sassano, and A. Astolfi. Dynamic scaling and observer design with application to adaptive control. *Automatica*, 45(12):2883–2889, 2009. 9

[83] S.M. Kay. *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Prentice Hall, 1993. 8

[84] N. Kazantzis and C. Kravaris. Nonlinear observer design using Lyapunov's auxiliary theorem. *Systems & Control Letters*, 34(5):241–247, 1998. 9

[85] H.K. Khalil. *Nonlinear systems*. Prentice Hall, 2nd edition, 1996. 53, 55, 77, 99, 100, 101, 104, 106, 107, 108, 114, 115, 116

[86] H.K. Khalil. *Nonlinear systems*. Prentice Hall, 3rd edition, 2000. 12, 31, 58, 69, 72, 94, 96, 98, 109, 111, 113

[87] H.K. Khalil. High-gain observers in nonlinear feedback control. In *Control, Automation and Systems, 2008. ICCAS 2008. International Conference on*, pages xlvii –lvii, 2008. 9, 10

[88] S.J. Kiebel, M.I. Garrido, R.J. Moran, and K.J. Friston. Dynamic causal modelling for EEG and MEG. *Cognitive neurodynamics*, 2(2):121–136, 2008. 9

[89] G. Kreisselmeier. Adaptive observers with exponential rate of convergence. *Automatic Control, IEEE Transactions on*, 22(1):2–8, 1977. 10, 84

[90] A.J. Krener and A. Isidori. Linearization by output injection and nonlinear observers. *Systems & Control Letters*, 3(1):47–52, 1983. 9, 30

[91] A.J. Krener and W. Respondek. Nonlinear observers with linearizable error dynamics. *SIAM Journal on Control and Optimization*, 23:197, 1985. 9, 30

[92] M. Krstic and A. Smyshlyaev. *Boundary control of PDEs: A course on backstepping designs*, volume 16. Society for Industrial Mathematics, 2008. 90

[93] L. Kuhlmann, D. Freestone, A. Lai, A.N. Burkitt, K. Fuller, D.B. Grayden, L. Seiderer, S. Vogrin, I.M.Y. Mareels, and M.J. Cook. Patient-specific bivariate-synchrony-based seizure prediction for short prediction horizons. *Epilepsy Research*, 91:214–231, 2010. 2

[94] H. Lei, J. Wei, and W. Lin. A global observer for observable autonomous systems with bounded solution trajectories. In *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC'05. 44th IEEE Conference on*, pages 1911–1916. IEEE, 2006. 9

[95] D. Liberzon. *Switching in systems and control*. Springer, 2003. 73, 79, 81, 121

[96] Y. Lin, E.D. Sontag, and Y. Wang. A Smooth Converse Lyapunov Theorem for Robust Stability. *SIAM Journal on Control and Optimization*, 34(1):124–160, 1996. 96

[97] F.L. Liu, M. Farza, M. M'saad, and H. Hammouri. Observer design for a class of uniformly observable MIMO nonlinear systems with coupled structure. In *Proceedings of the IFAC World Congress*, 2008. 9

[98] L. Ljung. *System identification: theory for the user*. Prentice Hall, Upper Saddle River, NJ, 1999. 35

[99] S.F.H. Lopes, A. Van Rotterdam, P. Barts, E. Van Heusden, and W. Burr. Models of neuronal populations: the basic mechanisms of rhythmicity. *Progress in brain research*, 45:281, 1976. 8, 18

[100] FH Lopes da Silva, A. Hoeks, H. Smits, and LH Zetterberg. Model of brain rhythmic activity. *Biological Cybernetics*, 15(1):27–37, 1974. 8, 18

[101] D. Luenberger. Observers for multivariable systems. *IEEE Transactions on Automatic Control*, 11(2):190 – 197, April 1966. 9

[102] D. Luenberger. An introduction to observers. *IEEE Transactions on Automatic Control*, 16(6):596 – 602, December 1971. 9

[103] Y. Mao, W. Tang, Ying Liu, and L. Kocarev. Identification of biological neurons using adaptive observers. *Cognitive Processing*, 10:41–53, 2009. 4

[104] R. Marine, G.L. Santosuosso, and P. Tomei. Robust adaptive observers for nonlinear systems with bounded disturbances. *IEEE Transactions on Automatic Control*, 46(6):967–972, 2001. 9, 10

[105] R. Marino and P. Tomei. Adaptive observers with arbitrary exponential rate of convergence for nonlinear systems. *IEEE Transactions on Automatic Control*, 40(7):1300 –1304, July 1995. 10

[106] R. Meier, H. Dittrich, A. Schulze-Bonhage, and A. Aertsen. Detecting epileptic seizures in long-term human EEG: a new approach to automatic online and real-time detection and classification of polymorphic seizure patterns. *Journal of Clinical Neurophysiology*, 25(3):119, 2008. 2

[107] P. Mirowski, D. Madhavan, Y. LeCun, and R. Kuzniecky. Classification of patterns of EEG synchronization for seizure prediction. *Clinical neurophysiology*, 120(11):1927–1940, 2009. 2

[108] P.E. Moraal and J.W. Grizzle. Observer design for nonlinear systems with discrete-time measurements. *IEEE Transactions on Automatic Control*, 40(3):395–404, 1995. 9

[109] F. Mormann, R.G. Andrzejak, C.E. Elger, and K. Lehnertz. Seizure prediction: the long and winding road. *Brain*, 130(2):314, 2007. 2

[110] A.S. Morse. Supervisory control of families of linear set-point controllers. 2. robustness. *IEEE Transactions on Automatic Control*, 42(11):1500–1515, 1997. 73

[111] V.B. Mountcastle. *Perceptual neuroscience: The cerebral cortex*. Harvard University Press, 1998. 6, 16

[112] K.S. Narendra and A.M. Annaswamy. *Stable adaptive systems*. Prentice-Hall, Inc., 1989. 79, 82, 84, 94

[113] H. Nijmeijer and T.I. Fossen. *New directions in nonlinear observer design*. Lecture notes in control and information sciences. Springer, 1999. 2, 9

[114] P.L. Nunez. The Brain Wave Equation: A Model for EEG. *Electroenchephalography and Clinical Neurophysiology*, 37:426, 1974. 90

[115] P.L. Nunez and R. Srinivasan. *Electric fields of the brain: the neurophysics of EEG*. Oxford University Press New York, 2006. 6, 14, 16

[116] I. Osorio, M.G. Frei, S. Sunderam, J. Giftakis, N.C. Bhavaraju, S.F. Schaffner, and S.B. Wilkinson. Automated seizure abatement in humans using electrical stimulation. *Annals of neurology*, 57(2):258–268, 2005. 91

[117] I. Osorio, M.G. Frei, and S.B. Wilkinson. Real-Time Automated Detection and Quantitative Analysis of Seizures and Short-Term Prediction of Clinical Onset. *Epilepsia*, 39(6):615–627, 1998. 2

[118] E. Panteley, A. Loría, and A. Teel. Relaxed persistency of excitation for uniform asymptotic stability. *IEEE Transactions on Automatic Control*, 46(12):1874–1886, 2001. 79, 82, 85, 94

[119] A.M. Pertew, H.J. Marquez, and Q. Zhao. H-infinity synthesis of unknown input observers for non-linear Lipschitz systems. *International Journal of Control*, 78(15):1155–1165, 2005. 9

[120] S. Raghavan and J.K. Hedrick. Observer design for a class of nonlinear systems. *International Journal of Control*, 59(2):515–528, 1994. 9, 37

[121] R. Rajamani. Observers for Lipschitz nonlinear systems. *IEEE Transactions on Automatic Control*, 43(3):397–401, 1998. 9, 10

[122] R. Rajamani and JK Hedrick. Adaptive observers for active automotive suspensions: theory and experiment. *IEEE Transactions on Control Systems Technology*, 3(1):86–93, 1995. 10

[123] M. Rodrigues, H. Hammouri, C. Mechmeche, and N.B. Braiek. A high gain observer based LMI approach. In *Proceedings of the IFAC World Congress*, 2008. 9, 30

[124] S.J. Schiff. *Neural Control Engineering: The Emerging Intersection Between Control Theory and Neuroscience*. Computational Neuroscience. The MIT Press, 2011. 2

[125] S.J. Schiff and T. Sauer. Kalman filter control of a model of spatiotemporal cortical dynamics. *Journal of Neural Engineering*, 5:1–8, 2008. 9

[126] Y. Shen and X. Xia. Semi-global finite-time observers for nonlinear systems. *Automatica*, 44(12):3152–3156, 2008. 9

[127] G.M. Shepherd. *The synaptic organization of the brain*. Oxford University Press New York, 2004. 6, 14

[128] H. Shim, J. H. Seo, and A. R. Teel. Nonlinear observer design via passivation of error dynamics. *Automatica*, 39(5):885 – 892, 2003. 9

[129] J.J. E. Slotine, J. K. Hedrick, and E. A. Misawa. On sliding observers for nonlinear systems. In *Proceedings of the American Control Conference*, pages 1794 –1800, 1986. 9

[130] B. Song and J.K. Hedrick. Observer-based dynamic surface control for a class of nonlinear systems: an lmi approach. *IEEE Transactions on Automatic Control*, 49(11):1995–2001, 2004. 9

[131] E.D. Sontag. Smooth stabilization implies coprime factorization. *IEEE Transactions on Automatic Control*, 34(4):435–443, 1989. 32, 53

[132] E.D. Sontag. *Mathematical control theory: deterministic finite dimensional systems*. Springer Verlag, 1998. 8

[133] E.D. Sontag. Input to state stability: Basic concepts and results. *Nonlinear and Optimal Control Theory*, 1932:163–220, 2008. 32, 105

[134] E.D. Sontag and A. Teel. Changing supply functions in input/state stable systems. *IEEE Transactions on Automatic Control*, 40(8):1476–1478, 1995. 102

[135] E.D. Sontag and Y. Wang. On characterizations of the input-to-state stability property. *Systems & Control Letters*, 24(5):351–359, 1995. 101

[136] R.J. Staba and A. Bragin. High-frequency oscillations and other electrophysiological biomarkers of epilepsy: underlying mechanisms. *Biomarkers*, 5(5):545–556, 2011. 3, 4

[137] W.C. Stacey and B. Litt. Technology insight: neuroengineering and epilepsydesigning devices for seizure control. *Nature Clinical Practice Neurology*, 4(4):190–201, 2008. 2

[138] C.J. Stam, J.P.M. Pijn, P. Suffczynski, and F.H. Lopes da Silva. Dynamics of the human alpha rhythm: evidence for non-linearity? *Clinical Neurophysiology*, 110(10):1801–1813, 1999. 8, 12, 18, 22, 35, 56, 129

[139] Ø.N. Stamnes, O.M. Aamo, and G.O. Kaasa. Redesign of adaptive observers for improved parameter identification in nonlinear systems. *Automatica*, 47(2):403–410, 2011. 10

[140] K.E. Stephan, L.M. Harrison, S.J. Kiebel, O. David, W.D. Penny, and K.J. Friston. Dynamic causal models of neural system dynamics: current state and future extensions. *Journal of biosciences*, 32(1):129–144, 2007. 9

[141] P. Suffczynski, S. Kalitzin, and FH Lopes Da Silva. Dynamics of non-convulsive epileptic phenomena modeled by a bistable neuronal network. *Neuroscience*, 126(2):467–484, 2004. 28

[142] F.E. Thau. Observing the state of non-linear dynamic systems. *International Journal of Control*, 17(3):471–479, 1973. 9

[143] I. Tokuda, U. Parlitz, L. Illing, M. Kennel, and H. Abarbanel. Parameter estimation for neuron models. *Experimental Chaos*, 676:251–256, 2003. 4

[144] Y. Totoki, K. Mitsunaga, H. Suemitsu, and T. Matsuo. Firing pattern estimation of synaptically coupled hindmarsh-rose neurons by adaptive observer. *Proceedings of the 18th international conference on Artificial Neural Networks*, 5164:338–347, 2008. 4

[145] R.D. Traub, A. Bibbig, F. LeBeau, E.H. Buhl, and M.A. Whittington. Cellular mechanisms of neuronal population oscillations in the hippocampus in vitro. *Annual Review of Neuroscience*, 27:247–278, 2004. 3

[146] J. Tsinias. Observer design for nonlinear systems. *Systems & Control Letters*, 13(2):135–142, 1989. 9

[147] J. Tsinias. Time-varying observers for a class of nonlinear systems. *Systems & Control Letters*, 57(12):1037 – 1047, 2008. 9

[148] I. Tyukin, E. Steur, H. Nijmeijer, D. Fairhurst, I. Song, A. Semyanov, and C. Van Leeuwen. State and parameter estimation for canonic models of neural oscillators. *International journal of neural systems*, 20(3):193, 2010. 4, 64

[149] I.Y. Tyukin, E. Steur, H. Nijmeijer, and C. van Leeuwen. State and parameter estimation for systems in non-canonical adaptive observer form. pages 6–11. 10

[150] G. Ullah and S. Schiff. Assimilating seizure dynamics. *PLoS computational biology*, 6(5):e1000776, 2010. 4, 9

[151] L. Vu, D. Chatterjee, and D. Liberzon. Input-to-state stability of switched systems and switching adaptive control. *Automatica*, 43:649–646, 2007. 75, 77, 80, 121

[152] L. Vu and D. Liberzon. Supervisory control of uncertain linear time-varying systems. *IEEE Transactions on Automatic Control*, 56(1):27–42, 2011. 10, 73, 75, 77, 78

[153] F. Wendling, F. Bartolomei, J.J. Bellanger, and P. Chauvel. Epileptic fast activity can be explained by a model of impaired GABAergic dendritic inhibition. *European Journal of Neuroscience*, 15(9):1499–1508, 2002. 4, 7, 12

[154] F. Wendling, J.J. Bellanger, F. Bartolomei, and P. Chauvel. Relevance of nonlinear lumped-parameter models in the anlysis of depth-eeg epileptic signals. *Biological Cybernetics*, 83:367–378, 2000. 66, 68

[155] F. Wendling, A. Hernandez, J.J. Bellanger, P. Chauvel, and F. Bartolomei. Interictal to ictal transition in human temporal lobe epilepsy: insights from a computational model of intracerebral EEG. *Journal of Clinical Neurophysiology*, 22(5):343, 2005. x, 4, 8, 16, 18, 19, 20, 22, 35, 37, 40, 41, 49, 56, 86, 91, 92, 93, 106, 131

[156] H.R. Wilson and J.D. Cowan. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Biological Cybernetics*, 13(2):55–80, 1973. 6, 8, 18

[157] F. Yang and R.W. Wilde. Observers for linear systems with unknown inputs. *IEEE Transactions on Automatic Control*, 33(7):677–681, 1988. 9

[158] A. Zemouche and M. Boutayeb. A unified H_∞ adaptive observer synthesis method for a class of systems with both Lipschitz and monotone nonlinearities. *Systems & Control Letters*, 58(4):282–288, 2009. 9, 10, 31, 50, 52, 59, 85, 86

[159] A. Zemouche, M. Boutayeb, and G.I. Bara. Observers for a class of Lipschitz systems with extension to H_∞ performance analysis. *Systems & Control Letters*, 57(1):18–27, 2008. 9

[160] Q.H. Zhang. Adaptive observer for multiple-input-multiple-output (MIMO) linear time-varying systems. *IEEE Transactions on Automatic Control*, 47(3):525–529, 2002. 10, 69, 111

[161] Q.H. Zhang. Revisiting different adaptive observers through a unified formulation. In *Proceedings of the IEEE Conference on Decision & Control and European Conference on Control*, pages 3067–3072, Seville, Spain, 2005. 9, 10, 64

[162] Q.H. Zhang and A. Clavel. Adaptive observer with exponential forgetting factor for linear time varying systems. In *Proceedings of the IEEE Conference on Decision and Control*, pages 3886–3891, 2001. 10, 69, 70, 111, 112

[163] F. Zhu and Z. Han. A note on observers for Lipschitz nonlinear systems. *IEEE Transactions on Automatic Control*, 47(10):1751–1754, 2002. 9

[164] M. Zijlmans, P. Jiruska, R. Zelmann, F.S. Leijten, J.G.R. Jefferys, and J. Gotman. High-frequency oscillations as a new biomarker in epilepsy. *Annals of neurology*, 71(2):169–178, 2012. 3

Author/s:
Chong, Michelle Siu Tze

Title:
Parameter and state estimation of nonlinear systems with applications in neuroscience

Date:
2013

Citation:
Chong, M. S. T. (2013). Parameter and state estimation of nonlinear systems with applications in neuroscience. PhD thesis, Department of Electrical and Electronic Engineering, The University of Melbourne.

Persistent Link:
http://hdl.handle.net/11343/38298